The Dissertation Committee for Brian Michael Ribnicky Certifies that this is the approved version of the following dissertation:

FACS: A High Throughput Method for Protein Export and Engineering

Committee:

George Georgiou, Supervisor

Brent Iverson

Brian Korgel

Charles Mullins

Charles Earhart

FACS: A High Throughput Method for Protein Export and Engineering

by

Brian Michael Ribnicky, B.E.; M.S.E

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2006

Dedication

To my parents for their never-ending love and support and to the rest of my family and friends for always brightening my day

FACS: A High Throughput Method for Protein Export and Engineering

Publication No._____

Brian Michael Ribnicky, PhD

The University of Texas at Austin, 2006

Supervisor: George Georgiou

Production of recombinant proteins is typically done in bacteria, i.e. *E. coli*, since they can be grown quickly, inexpensively and can be used to express a variety of proteins. However, obstacles, such as the presence of cytoplasmic proteases and low periplasmic protein yields, exist that prevent *E. coli* from being more widely used for protein production. By using the methods of protein engineering coupled with high-throughput screening by Fluorescence Activated Cell Sorting (FACS), recombinant protein yields in *E. coli* may be increased which would lead to the greater utilization of *E. coli* as a recombinant protein production host.

The recently discovered Twin Arginine Translocation (Tat) system allows the export of fully-folded proteins across the inner cell membrane of *E. coli*. The Tat system now allows protein engineers to take advantage of the available protein folding machinery in the cytoplasm, expanding the range of recombinant proteins that can be produced in bacteria. However, since the efficiency and yield of export via Tat is low,

we sought to examine whether mutations within the mature part of the protein can result in increased yield of periplasmic proteins. Using directed evolution and FACS, the export of the anti-digoxin 26-10 scFv was increased in an oxidizing strain of *E. coli*. One isolated clone showed increased export of 26-10 to the periplasm. *In vitro* analysis of the thermodynamic stability and folding kinetics indicated that the mutant scFv exhibits faster folding kinetics, and this is likely to be responsible for the improved export to the periplasmic space.

In addition, in order further increase the export of the 26-10 scFv through the Tat pathway, we examined the effect of the co-expression of certain chaperone and Tat translocon proteins that were previously shown to partially relieve the saturation of the Tat pathway. Co-expression of the Tat translocon proteins, TatA, TatB and TatC, resulted in an additional 2.4-fold increase in cell fluorescence yielding an overall cell fluorescence increase of 5-fold as compared to the parental 26-10 scFv. The effect of media, pH and temperature and the use of different oxidizing strains of *E. coli* on cell fluorescence and growth were also investigated but did not yield any improvement on cell fluorescence.

Finally, flow cytometry was utilized in an attempt to develop a genetic system for the better understanding of the Clp system and for studying how amino acid substitutions might affect the mechanical stability of protein substrates. The Clp system recognizes proteins tagged for degradation, unfolds them by applying a mechanical force to the protein and degrades them into smaller peptides. A FACS screen was employed to screen a genomic library for genes that can act as negative regulators of ClpXP degradation. Unfortunately, no such genes could be isolated. In addition, since the system applies a mechanical force to unfold proteins, the Clp system could be used as an assay to engineer more stable proteins against mechanical denaturation. Using GFP

tagged with the SsrA peptide as a fluorescent reporter in a FACS assay, the screening of several mutant libraries indicated that no mutations could stabilize GFP against degradation by ClpXP.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction, Background and State of the Art

1.1 INTRODUCTION

Protein engineering is the application of science and engineering to improve existing proteins and enzymes. Two different, but complementary, strategies are used in protein engineering. Rational protein design takes advantage of the knowledge of a protein's structure and function to make directed amino acid changes in order to improve certain properties of the protein. However, since the detailed structure of a protein may not be known, it is difficult to predict the location and changes necessary to obtain the desired results. The second approach to protein engineering, which does not require detailed structural or functional information, includes several strategies collectively termed directed evolution. In the most common application of directed evolution, random mutagenesis is applied to the whole protein, and a selection method is used that can result in the isolation of the individual mutant proteins with the desired properties. Multiple rounds of mutation and selection can be utilized to further enhance the properties of the isolated clones. However, directed evolution requires high-throughput screening (HTS) methods in order to ensure that a large number of combinations of random mutations are investigated. Thus, in any directed evolution experiment, the design of a high-throughput screen is one of the most critical considerations.

1.1.1 Protein Engineering and Antibodies

The capabilities of protein engineering have been increasing as more knowledge of protein structure and function becomes available along with advancements in high-throughput screening. The first successes of protein engineering improved a protein's physical properties to withstand varying environmental conditions such as pH and temperature [1]. More recently, researchers have been able to engineer proteins with desirable therapeutic properties for the specific delivery of drugs to increase their potency and reduce possible side effects [2-4]. More importantly, a special class of proteins known as antibodies can be engineered to target biologically important molecules like hormones or their receptors [5, 6]. There are currently 18 antibodies approved for therapeutic use in the United States [7] with another 150 in clinical development [8].

An antibody, or immunoglobulin, is produced by the immune system to recognize specific antigens such as bacteria, viruses, and infected cells. Upon binding to a unique target, the antibody tags the microbe or infected cell for clearance by other parts of the immune system. In addition, targets can also be neutralized when the binding antibody blocks an essential part of the microorganism necessary for its survival or invasion of the host cell. The most prevalent immunoglobulin in human serum is immunoglobulin G (IgG). It consists of four polypeptide chains, 2 identical light chains and 2 identical heavy chains in a characteristic 'Y' shape (see Figure 1.1, [7]). A disulfide bond links each half of the 'Y'. The Fc region of the IgG (C_{H2} and C_{H3}) is also known as the constant fragment and is responsible for the effector functions (i.e. macrophage recognition and complement activation) of the IgG during an immune response. The antigen binding fragment or Fab makes up each of the arms of the 'Y' shape and contains one set of variable chains (V_H and V_L) and their associated constant chains (C_{H1} and $C_{\lambda(\kappa)}$). The variable fragment (Fv) region on the end of each Fab consists of the V_H and

V_L chains, and each chain has three hypervariable loops known as complementary determining regions (CDRs). The polypeptide regions surrounding the CDRs are known as the framework regions, and along with the CDRs, play a role in the antigen binding properties of the antibody. By targeting the CDRs and framework regions at the nucleic acid level, the antigen binding properties can be altered to increase an antibody's specificity or affinity.

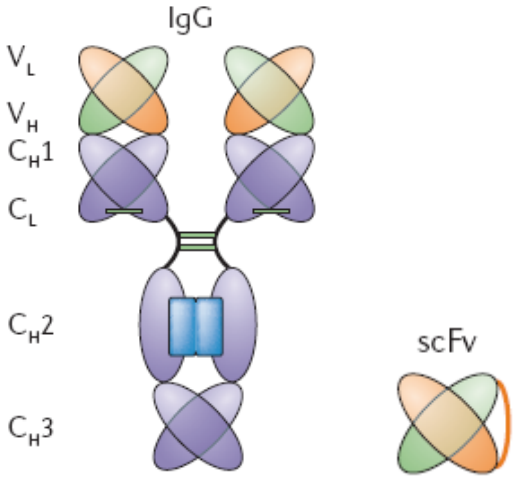

Figure 1.1. Diagram of an IgG and scFv antibody. Single-chain antibody fragments are comprised of the variable domains from the heavy (V_H) and light chain (V_L) of the IgG joined by a peptide linker of about 15 amino acids. Constant domains (C) of both the heavy and light chains are linked by disulfide bonds shown as green lines. Figure adapted from [7].

Since the binding domain of the antibody is comprised of the V_H and V_L chains, it is only necessary to manipulate these two chains to alter the specificity or affinity of the antibody. Therefore, it was these two regions that were first amplified and expressed in microbial hosts [9]. In order to produce a single-chain antibody fragment (scFv), a flexible peptide linker was added between the two variable domains [10]. The scFv represents one entire binding site of the antibody molecule and can be easily expressed in bacteria. The small fragment size (~30 kDa compared to ~150 kDa for IgG) is easier to express and produce in microbial cells which lowers production time and costs. In

addition, since the affinity of scFvs is similar to each of the binding sites on an IgG molecule [11], the scFv can be used in conjunction with recombinant DNA techniques in bacteria as a means to engineer improved antibodies. Several antibody fragments are in clinical development for multiple diseases including cancer [12]. However, product yields of antibody fragments in *Escherichia coli (E. coli)* have been inadequate, and therefore protein engineering can play a vital role in increasing the expression and overall production of antibodies from bacteria.

1.2 HIGH-THROUGHPUT SCREENING AND PROTEIN ENGINEERING

1.2.1 Directed Evolution

A combinatorial library is a collection of cloned DNA genes or proteins that have been mutated such that different combinations of amino acids are represented. These libraries can be constructed by various methods including error-prone polymerase chain reaction (PCR) in which the whole or part of a gene is mutated [13, 14]. Directed evolution relies on the creation of these libraries which can then be screened for "gain of function" mutants towards a specific application [15]. Directed evolution is typically an iterative approach: the gene encoding the protein of interest is mutated to create the library of gene variants, the library is screened to isolate the clones that posses the desired properties and the isolated clones can then be subjected to additional rounds of mutation and selection.

In order for directed evolution to be successful, mutations at the DNA level have to be linked to changes in the protein's function. In other words, the genotype (DNA level) must be linked to the phenotype (i.e. a function mediated by the protein), and this linkage must be maintained through the selection or screening process. Since any combinatorial approach relies on methods for the rapid measurement of the phenotypes of the mutant library to identify the clones with improved properties [16], the linkage method is typically developed with the selection system in mind.

Screening individual clones on microtiter well plates or agar plates limits the amount of clones that can be examined effectively to typically 10^4 [17]. HTS methods which typical employ fluorescence for detection of the desired function can be employed to interrogate a larger number of cells producing different mutant proteins. For example, using HTS and a fluorescent phenotype, directed evolution has allowed for the engineering of enzymes with increased substrate affinity [18, 19].

1.2.2 Microbial Cell Display

One method to establish a link between genotype and phenotype is microbial cell display. This approach involves the expression of protein libraries via plasmid expression vectors in a suitable host organism, typically bacteria or yeast. However, the location of protein expression in Gram-negative bacteria, such as *E. coli*, can determine the type of display system used.

The periplasmic space of *E. coli*, located between the lipopolysaccharide outer membrane and the inner membrane, has an oxidizing environment which allows for the formation of disulfide bonds (see Figure 1.2, [20]). Proteins that contain disulfide bonds are therefore typically expressed in the periplasm to allow the protein to fold correctly. The cytoplasm, on the other hand, has a reducing environment and contains the majority of the cell's folding machinery [21] such as molecular chaperones that prevent protein aggregation [22]. Hence, cytoplasmic expression may allow these chaperones to help proteins fold that are normally prone to aggregation.

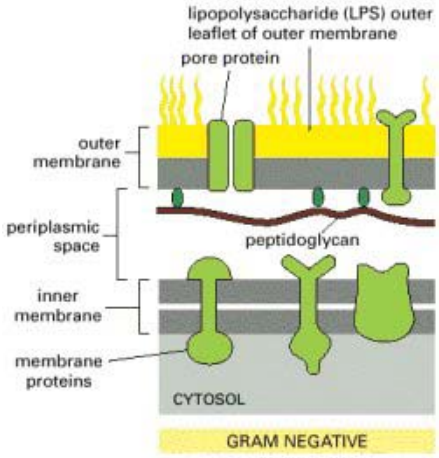

Figure 1.2. The cell envelop of Gram negative bacteria. The inner membrane separates the cytoplasm from the periplasmic space, while the outer membrane separates the periplasmic space from the extracellular environment. The inner membrane is a phospholipid bilayer while the outer membrane is comprised of a phospholipid leaflet and a lipopolysaccharide leaflet. Also shown is the peptidoglycan wall connected to the outer membrane by lipoproteins and embedded proteins in both membranes. Figure adapted from [20].

In *E. coli* display systems, the proteins are usually expressed in the periplasm to allow for correct folding. Secreted proteins can be anchored to the periplasmic side of the inner membrane. Typically, when screening antibody libraries displayed in bacteria or yeast, the cells are incubated with fluorescently labeled antigen to allow selection based on the fluorescence level. Therefore, cells isolated based on high fluorescence may indicate an increase in the affinity of the antibody or an increase in the number of antibodies displayed by each cell. This results in the ability to select for mutant antibodies with better expression levels, export, or folding properties or increased amounts of soluble protein.

The high-throughput analysis and selection of antibodies exhibiting higher fluorescence is typically carried out using a method known as flow cytometry (described in detail in a later section) [23]. More specifically, flow cytometry has been used to sort and isolate high-affinity antibodies against a target antigen [23-25].

However, most fluorescent ligands cannot enter microbial cells so the protein of interest must be accessible to the extracellular fluid. In order to accomplish this, the Periplasmic Expression with Cytometric Screening (PECS) [25] system relies on the export of the protein into the periplasm of *E. coli* and permeabilization of the outer membrane to allow access by fluorescent ligands (see Figure 1.3). This allows fluorescent molecules to enter the periplasm without the leakage of proteins from the periplasm or a loss in cell viability [25]. Since the fluorescent marker enters from the extracellular space, no fusion partners are needed to tag the protein. In addition, only soluble proteins will be selected since insoluble protein will not bind the fluorescent ligand and be selected, which is useful for purification and characterization. Thus, the PECS system allows the screening of large protein libraries while providing the necessary link between genotype and phenotype.

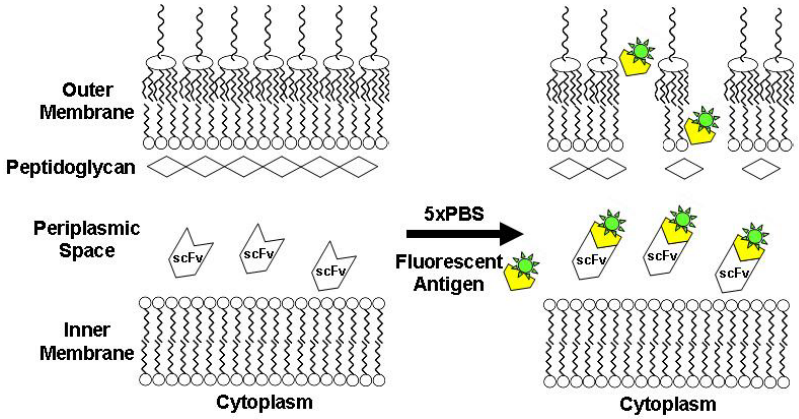

Figure 1.3. PECS diagram. Cells expressing the protein of interest shown as single chain antibody fragments (scFv) are incubated in hypertonic buffer (i.e. buffer of high osmolarity such as 5xPBS (phosphate buffered saline)) to permeabilize the outer membrane to the fluorescently labeled antigen (green symbols). The cells are then washed to remove unbound antigen, and cells are selected on their level of fluorescence.

1.2.3 Flow Cytometry (FC)

Flow cytometry is able to analyze multiple properties of single cells in a large population quickly and it has become widely used as a HTS method [26]. Flow cytometry can measure phenotypic and molecular properties of individual cells using fluorescent probes or markers which bind to specific molecules in or on the cell [27], and it has been used to simultaneously measure 13 different cellular parameters [28]. Cells are forced to flow single file in front of a laser beam. As the cells flow by, the laser excites the fluorescent molecules, and the emitted light is collected by detectors which then translate the light into a proportional electronic signal. In addition, the diffracted light or forward scatter is related to the cell surface area while reflected and refracted light or side-scatter yields information about cell granularity (see Figure 1.4, [27]).

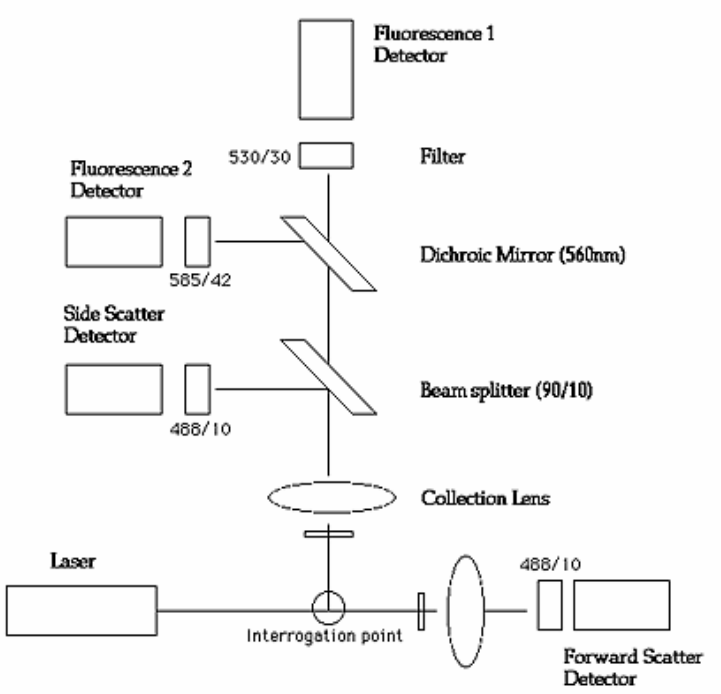

Figure 1.4. Simplified diagram of a flow cytometer. As shown, a laser is focused at the interrogation point at which cells pass through individually. Properties of the cell can be determined by the light collected by the side scatter and forward scatter detectors and by the fluorescence detectors. Different filters can be used depending upon the excitation and emission spectra of the fluorescent probes used. Figure adapted from [27].

Fluorescence Activated Cell Sorting (FACS) is routinely used in research settings to conduct immunophenotyping, cell counting, DNA analysis and isolation of antibody fragments that bind target antigens by sorting [23, 29]. During FACS, a population of cells is defined by its forward and side-scatter characteristics, and then individual cells can be collected from the population based on their fluorescence. When a cell's properties match the sort criteria, the FACS system moves a catcher tube into the fluid stream to capture that single cell (see Figure 1.5, [30]). The catcher tube can move in and out of the cell stream approximately 300 times per second, and the FACS system can sort over 10^7 events per hour [31]. It has been shown that even at these high sort rates, rare clones with frequencies as low as 10^6 to 10^7 can be isolated from heterogeneous

populations [32, 33]. Due to the quantitative information collected about each cell in the library, FACS can provide a means to optimize library construction and sorting conditions to better obtain clones of interest [34, 35]. Thus, this high-throughput screening method coupled with the quantitative analysis of flow cytometry has made FACS a useful tool for isolating mutants of interest from protein libraries.

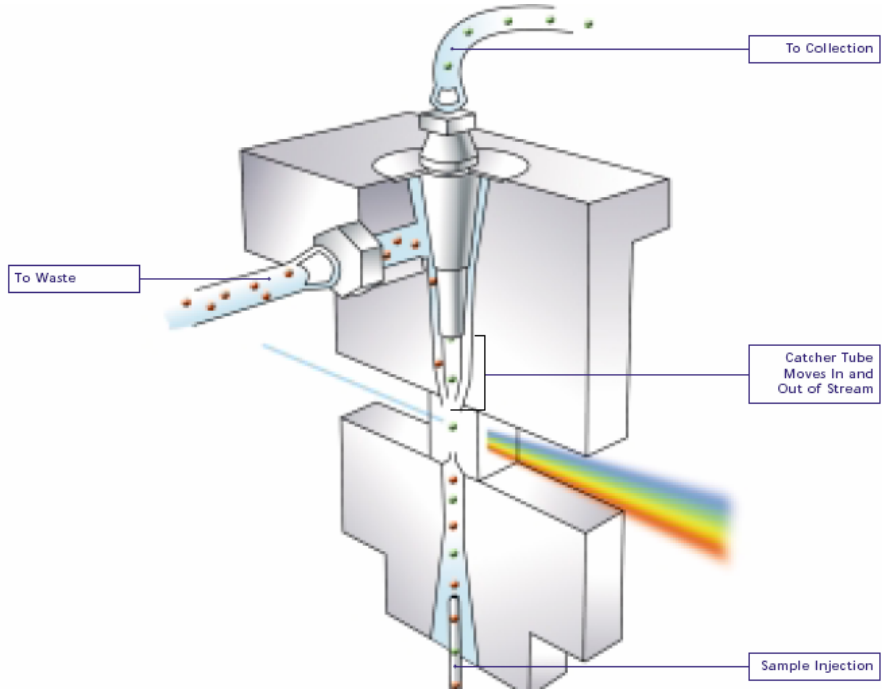

Figure 1.5. Schematic of cell sorting system. As the individual cells pass the laser, its fluorescence is detected and compared with the set sort conditions. If the sort conditions are satisfied, a catcher tube will move into the stream to collect the cell. Otherwise, the cell will continue to the waste tank. Figure adapted from [30].

1.3 PROTEIN ENGINEERING AND PRODUCTION

The current biopharmaceutical protein market generates over $30 billion per year in revenue and is expected to double by the end of the decade [36, 37]. The majority of recombinant proteins are currently produced in bacteria due the lower cost, higher yields, and ease of growth in the laboratory [21, 38, 39]. Specifically, *E. coli* bacteria are used for about 40% of the total recombinant protein production [36].

However, obstacles exist that prevent *E. coli* from being more widely used for protein production. Although cytoplasmic protein expression allows the protein to take advantage of the molecular chaperones, cofactors, and foldases present in the cytoplasm to ensure proper folding, the cytoplasm also contains many protease systems. One such system, the *E. coli* Clp system (discussed below), can reduce protein yields through degradation. Being able to control proteases such as those in the Clp system would reduce cytoplasmic degradation of recombinant proteins.

The export of proteins into the periplasm may improve yields and also reduce further downstream processing that would normally be required to remove cytoplasmic components. Since the periplasm lacks many of the chaperones of the cytoplasm, exporting fully folded proteins to the periplasm is desirable. In bacteria, the Twin Arginine Translocation (Tat) pathway (discussed in detail in a later section) is the only known system that functions to transport folded proteins across the cytoplasmic membrane. However, this process is relatively inefficient, resulting in low protein yields particularly when compared to the export pathways that transport polypeptides in an unfolded conformation. Nevertheless, protein engineering can be utilized to increase the efficiency of export by increasing the export competence of the protein. In addition, engineering the export machinery or new strains of *E. coli* may also increase export flux. Therefore, protein engineering coupled with HTS can be used to increase recombinant

protein yields in *E. coli* and lead to greater utilization of *E. coli* as a recombinant protein production host.

1.4 BACKGROUND

1.4.1 Protein Export in *E. coli*

For protein production purposes, it is often necessary to transport the proteins out of the cytoplasm after they are synthesized [40] since cytoplasmic expression leads to additional downstream processing steps. For instance, cytoplasmic proteins must be harvested after disruption of both cell walls and separated from all other unwanted cytoplasmic components whereas disruption of the outer membrane and collection of periplasmic proteins is less complicated [41]. In addition, protein aggregation in the cytoplasm can occur when overexpressing a heterologous protein [42, 43] which requires additional steps to refold the protein to obtain its biologically active form [36]. Finally, the total product yield may be affected by cytoplasmic protease degradation.

For the above reasons, the expression of proteins in secreted form is highly desirable for biotechnology purposes. Bacteria contain multiple protein export pathways. The two most important pathways, designated as Sec and Tat are similar in that they 1) use a signal peptide or N-terminal tag to target the protein to the pathway translocon, 2) use an energy source to translocate the protein through the pore, and 3) maintain the integrity and function of the inner membrane while allowing large proteins to be transported. Both of these pathways are briefly discussed in the following sections.

1.4.1.A. The Sec Pathway

In bacteria, the major pathway of protein translocation from the cytoplasm to the periplasmic space is through the General Secretory (Sec) translocon (see Figure 1.6, [44]). It is comprised of a motor domain, the ATPase of the SecA protein, and a membrane embedded channel consisting of proteins SecY, SecE and SecG [45]. Polypeptides are secreted in a largely unfolded form, threading their way through the

SecYEG pore by SecA. A Sec-specific chaperone, SecB recognizes the Sec signal sequence and targets the preprotein to the Sec pathway by its ability to bind SecA and maintain the polypeptide in an unfolded state [46]. SecA then feeds the polypeptide into the translocase through a repetitive ratchet-like motion [47], with approximately 20 amino acids inserted into the pore per cycle of ATP until the polypeptide has been completely secreted [48, 49]. In the periplasm, the signal peptide is cleaved by signal peptidase I, and the protein folds on its own or with the help of periplasmic oxidases or isomerases.

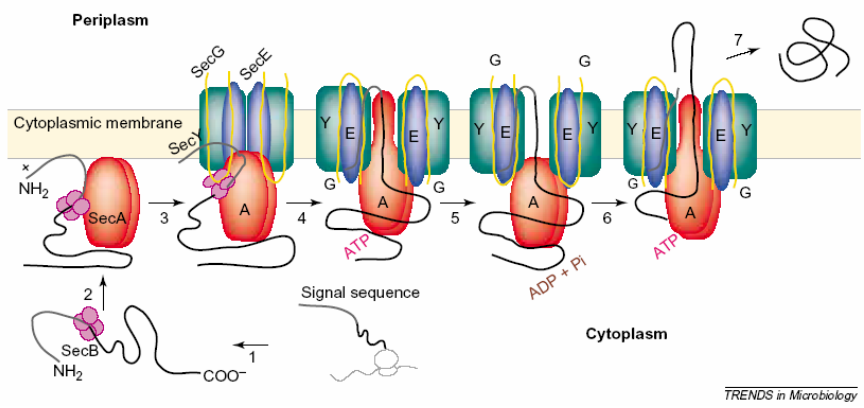

Figure 1.6. Model of the Sec pathway. As shown, SecB recognizes the signal sequence and targets the polypeptide to SecA which then translocates the polypeptide through the SecYEG pore by expending ATP. Figure adapted from [44].

Although the Sec pathway is used extensively for protein expression, it has some limitations. When incompatible heterologous polypeptides insert into the SecYEG pore, they become stuck and are unable to translocate [50]. This effectively blocks the export of all native Sec proteins causing cell lysis. In addition, some heterologous secreted proteins are toxic to the cell and also misfold in the periplasm rendering them inactive [51]. This is due to the fact that the periplasm lacks most of the ATP-dependent chaperones associated with the cytoplasm [52]. Finally, many enzymes require cofactors which need to be incorporated into the protein during folding. Since the Sec pore is

narrow, the system can only transport unfolded proteins [53] so cofactor insertion cannot occur and therefore limits the types of recombinant enzymes which can be secreted by the Sec system [54].

1.4.1.B. The Tat Pathway

The recently discovered Tat pathway is able to translocate fully folded proteins in their compact, native-like state across the inner membrane [55, 56]. With its discovery nearly ten years ago and the early evidence suggesting the pathway exports folded proteins [57], it appeared likely that the Tat pathway may be useful for potential engineering and production of proteins that cannot be transported across the inner membrane in an active form via the Sec pathway. For example, green fluorescent protein (GFP) is nonfluorescent when exported to the periplasm through the Sec pathway [58], but GFP fusions that are exported through the Tat pathway have been shown to be fluorescent and capable of being used as molecular tags [59].

Although only responsible for about 5-8% of secreted proteins in *E. coli* [60], most of these proteins are considered to require cofactor insertion [61], and the majority are cofactor-containing redox proteins [62, 63]. These cofactors are synthesized in the cytoplasm and are integrated into the protein during folding [63], making them incompatible with Sec transport. Also, proteins requiring the assembly of multiple subunits in the cytoplasm can be exported by Tat [55]. The Tat pathway has been shown to export proteins with molecular weights up to 120 kDa [57].

The Tat pathway is named for the invariant arginine residues in the N-terminal consensus motif, SRRXFLK, where X is any amino acid, of the Tat signal peptide [64]. There are three domains in a signal peptide: a positively charged N-terminal domain, a hydrophobic domain, and the C-terminal domain. The signal peptide is necessary to initiate export across the membrane in the majority of secretion systems [56]. Tat signal

peptides are approximately 14 amino acids longer, are significantly less hydrophobic in the middle of the signal, and are more positively charged at the end of the signal than Sec signal peptides [65, 66].

The Tat operon is comprised of five genes, *tatA*, *tatB*, *tatC*, *tatD* and *tatE*, but TatD has been shown to play no role in Tat pathway export [67], and *tatE* is a duplication of *tatA* [68]. Only the TatA, TatB and TatC proteins are necessary for Tat export, and these proteins contain at least one transmembrane helix which is a hydrophobic alpha helix that is inserted across the phospholipid bilayer of the inner membrane. TatA and TatB are very similar in sequence and structure in that they contain an N-terminal transmembrane helix followed by a short hinge region and then an amphipathic helix [69] (see Figure 1.7, [70]). Studies have shown that the TatB amphipathic helix is necessary for Tat transport [71], and that it contacts the entire signal peptide and some of the mature protein *in vitro* [72]. The transmembrane helix of TatA is required for interactions with TatB and export [71, 73]. TatC has six transmembrane helices and both the N- and C-termini on the cytoplasmic of the inner membrane [74, 75] indicating that the cytoplasmic side of TatC is important for function. It has been shown that signal peptides interact with TatC and that this interaction requires the twin-arginine motif of the Tat signal peptide *in vitro* [72, 76].

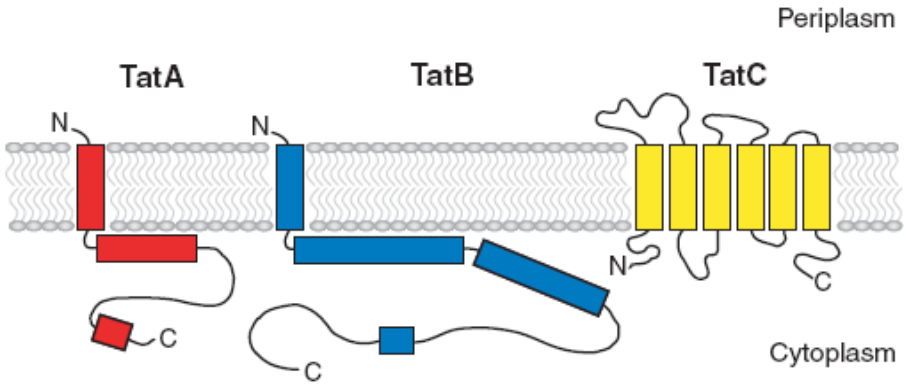

Figure 1.7. Predicted structure and topology of the *E. coli* Tat components. Boxes represent helical regions. Figure adapted from [70].

The current model of Tat translocation proposes that the Tat proteins remain as separate complexes and only come together to form a complete pore after substrate binding [72, 77]. This allows the formation of different size pores to match the folded Tat substrates which restricts the possibility of ion leakage through the pore while the substrate is being translocated (see Figure 1.8, [70]). TatA is expressed by cells approximately 20 times more than TatB or TatC [78], indicating that TatA most likely forms the majority of the Tat pore [77, 79]. In addition, electron microscopy of a Tat complex showed pore-like structures comprised mainly of TatA that were blocked on the cytoplasmic side of the inner membrane [80]. Studies have shown that TatB and TatC form membrane complexes with multiple copies of each protein in equimolar amounts [81]. In addition, membrane complexes of TatA and TatB have also been purified along with complexes containing TatA, TatB, and TatC in which TatA is present in a large molar excess [82]. Most likely, the Tat signal peptide is recognized by Tat-specific chaperones which target the pre-protein to TatBC after folding and cofactor insertion which results in the recruitment of TatA to form the protein transport channel (see Figure 1.8).

Transport through the Tat pathway is driven by the proton motive force across the membrane rather than ATP hydrolysis as in the Sec system [83]. The proton motive force is the sum of the electrical gradient due to the charge differential that exists across the membrane, and the chemical gradient caused by concentration differences on each side of the membrane [84]. The Sec system requires one ATP molecule per 20 amino acid residues that are translocated [49] while it has been calculated that the Tat pathway uses the energy equivalent of approximately 10,000 ATP molecules per protein [85]. Thus, the Tat pathway is more energetically demanding which may be the reason so few proteins are Tat substrates.

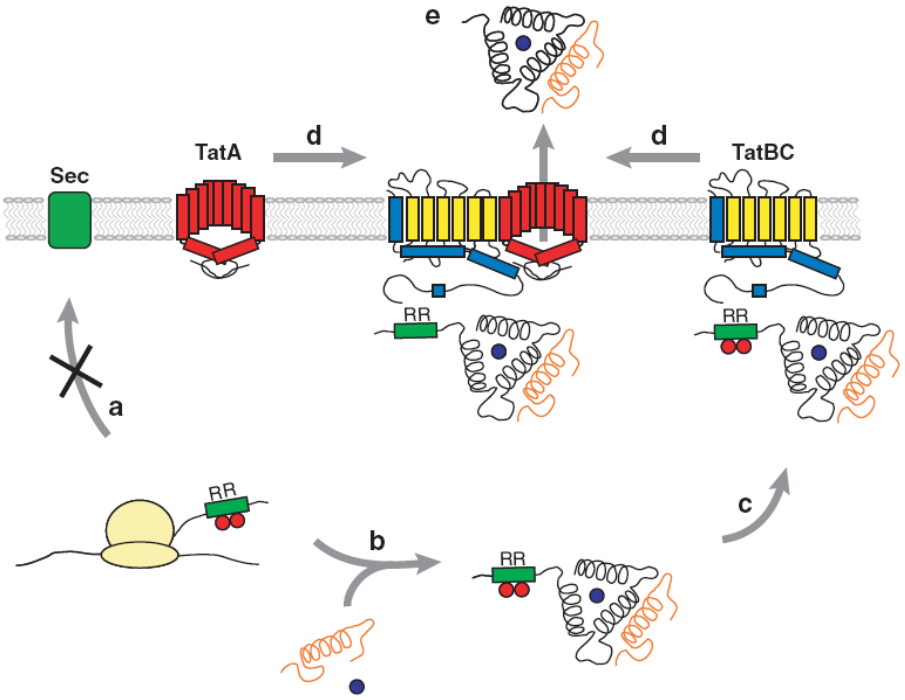

Figure 1.8. Model of Tat targeting and transport. After translation, the signal peptide and binding of Tat-specific chaperones (maroon circles) aid the pre-protein to avoid being targeted to other pathways (a) such as Sec. After folding and any necessary cofactor insertion or subunit association (b), the pre-protein is targeted to the TatBC complex (c). The pore is formed by TatA, and translocation is driven by the proton motive force (d). On the periplasmic side of the membrane, the signal peptide is removed, and the mature protein is released (e). TatA, TatB, and TatC are shown in maroon, blue, and gold, respectively. Figure adapted from [70].

1.4.2 ClpXP and ClpAP Proteases

Proteolysis plays a major role in the regulation of metabolic processes such as cell cycle control [86] and a cell's response to environmental factors and DNA damage [87, 88]. It is an irreversible mechanism that is energy dependent, and it is the main process by which cells degrade proteins that have spontaneously denatured, contain translation errors in the peptide chain or have accumulated mutations that interfere with the protein's function [89, 90]. Also, by degrading key rate-limiting enzymes, a cell's metabolic activities can be regulated [91]. Proteases are extremely important during stress responses to environmental factors during which many proteins become damaged [92]. Therefore, proteases are typically induced during these stress responses to prevent the accumulation of abnormal protein. In prokaryotes, the proteases ClpXP, ClpAP, Lon, HslUV and FtsH carry out the degradation of target proteins [93].

The serine-type ClpP protease from *E. coli* is one of the best characterized proteases, and the Clp (caseinolytic protease) proteins are well conserved throughout most prokaryotes and in some organelles of eukaryotes [94]. ClpA, an ATP-dependent chaperone, was the first protein of the Clp family to be isolated and was found to degrade casein when complexed with ClpP [95, 96]. ClpX was later identified as another ATP-dependent chaperone similar to ClpA [97]. ClpP can only degrade small peptides as folded proteins must first be unfolded through ATP hydrolysis by ClpA or ClpX [98, 99].

1.4.2.A. The ClpP Serine Protease

For both proteases, ClpP is the enzyme whose active proteolytic sites are within an internal cavity formed by two seven-subunit rings comprised of 14 identical subunits and are only accessible through narrow axial channels of approximately 50 angstroms in diameter [100-102] (see Figure 1.9A&B [89, 103]) . Each of these subunits contain a catalytic triad composed of serine-histidine-aspartic acid residues and is only accessible

to small peptides or larger, unfolded proteins translocated by either chaperone, ClpX or ClpA [102]. The *E. coli clpP* gene encodes a protein of 207 amino acids although the first 14 residues are autolytically cleaved to yield a mature protein of 21.5 kDa [96]. This self-processing may serve as a regulatory mechanism since only the mature protein is active [104]. Recent work has shown that the N-terminal region of the ClpP subunits form flexible loops that may help it associate with either chaperone [105].

1.4.2.B. The ClpA and ClpX Chaperones

In order to obtain full proteolytic activity, ClpP must associate with one of the two related ATPase subunits, ClpA or ClpX, both members of the AAA+ (ATPases Associated with diverse cellular Activities) superfamily of proteins. ClpA and ClpX are both ATP-dependent chaperones which are composed of six subunits arranged in a hexagonal ring that can associate with either or both sides of ClpP, but recognize different degradation tags [104, 105] (see Figure 1.9C [103]). ClpA and ClpX have the role of binding the target protein by recognizing a degradation tag [104], unfolding the protein and translocating the protein into the proteolytic chamber of ClpP [106, 107].

The two chaperones are very similar except that ClpA (84 kDa) has two ATPase domains while ClpX (46 kDa) only has one [97]. ClpX has been more widely studied and is comprised of three domains: an N-terminal domain containing a zinc binding domain [106], a central ATPase domain (or two for ClpA), and a C-terminal sensor and substrate discrimination (SSD) domain. Although this SSD domain can interact with some substrates [107], it is more likely that the N-terminal domain is responsible for substrate specificity since this domain is similar to the I-domain of *E. coli* HslU which has been shown to be responsible for substrate specificity [108]. Studies have also shown that ClpX activity is not affected when the N-terminal domain is removed [109]. The ClpX C-terminal domain has been shown to contain a conserved isoleucine-glycine-

phenylalanine (IGF) tripeptide which is located on an extended loop. Six tripeptides from the ClpX hexamer make hydrophobic contacts with the ClpP heptmer [110] (see Figure 1.10 [111]). Hexamerization occurs in the presence of ATP, Zn^{++} and Mg^{++} [106, 109, 112] and forms the axial pore through which substrate proteins are threaded, unfolded and translocated into the pore of ClpP for degradation [113]. In addition to degradation, ClpA and ClpX can also act as independent chaperones in which they bind an unfolded protein to allow remodeling to occur such that upon release, the protein folds into its correct confirmation [114] (see Figure 1.11 [103]).

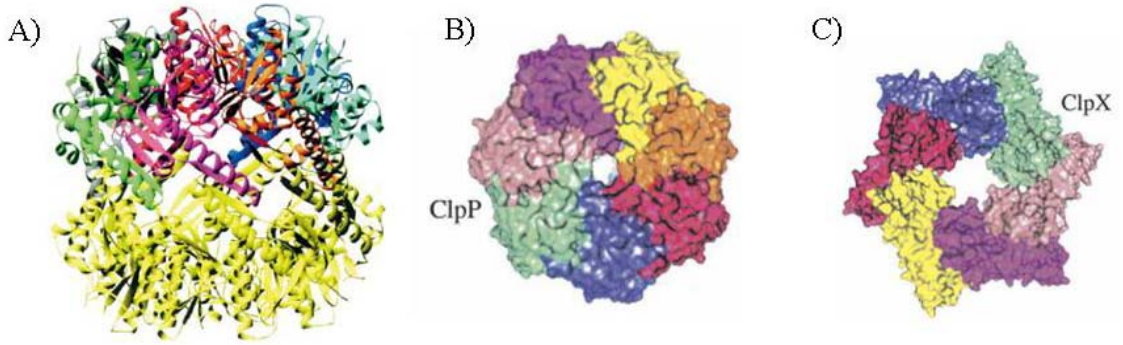

Figure 1.9. ClpP and ClpX representations. A) Ribbon model of ClpP showing the two stacked heptameric rings; the bottom in yellow and the top as colored subunits. B) Spacefill model of ClpP looking down through the axial pore. C) Spacefill model of ClpX also looking down through the axial pore. Both spacefill models show each subunit as a different color. Figure adapted from [89, 103].

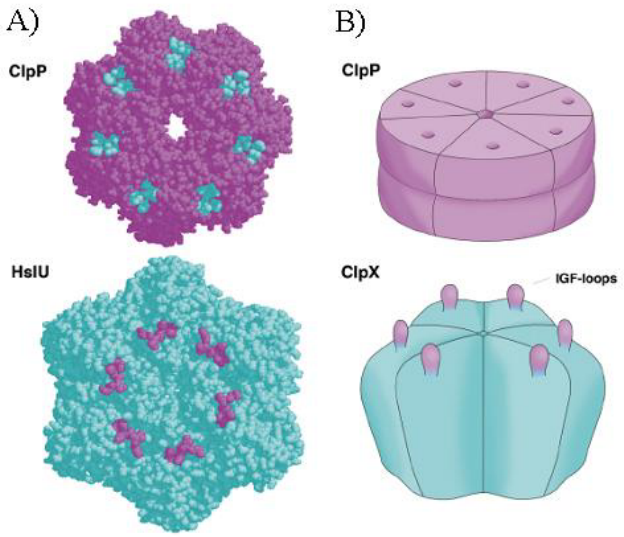

Figure 1.10. Model for ClpP and ClpX interaction. A) Spacefill model of ClpP and HslU indicating the possible interaction points. HslU is an analog of ClpX and does not bind ClpP. B) Cartoon showing how the hydrophobic pockets of ClpP may interact with the IGF loops of ClpX. Figure adapted from [111].

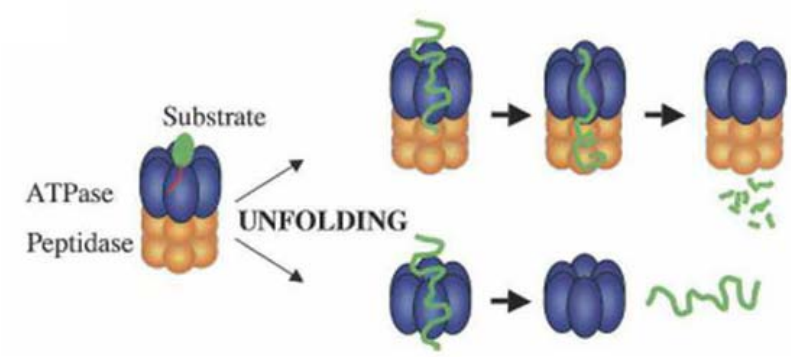

Figure 1.11. Schematic showing the dual roles of the ATPase. The ATPases, ClpX and ClpP are not only involved in substrate degradation. They can also bind unfolded or incorrectly folded substrates, unfold them and release them so that they may attempt to fold correctly again. Figure adapted from [103].

1.4.2.C. Substrate Recognition

Substrate recognition occurs through binding of either N- or C-terminal motifs intrinsic to the protein or through tags acquired after or during translation such as the SsrA protein tag [115]. These tags include three N-terminal and two C-terminal motifs identified from fifty ClpXP substrates [116] which include the plasmid P1 RepA and MuA transposase [117]. The motifs are comprised mostly of hydrophobic residues although one N-terminal motif contains polar residues. Transferring these motifs to other proteins makes them substrate for hydrolysis by ClpX [116]. The motifs can be moved to other positions in the protein, but their placement at the either end of the protein is more favorable due to the accessibility of the ends to ClpX [118]. Other efforts have shown that some tags recognized by ClpX can also be recognized by ClpA in certain cases [119, 120].

The SsrA tag of *E. coli* has the sequence of AANDENYALAA and is a hydrophobic tag that is grafted to proteins that are stalled in the ribosome during translation [121] (see Figure 1.12 [122]). It is translated onto a protein by an RNA molecule that has both transfer RNA activity and an mRNA sequence. When a ribosome stalls on a mRNA either due to amino acid starvation, rare codon usage or a lack of a stop codon [123], the alanine-charged-SsrA RNA enters the A site of the ribosome, causing the stalled polypeptide to be transpeptidated to the alanine and transferred to the SsrA RNA. The mRNA in the SsrA is then translated by the ribosome, adding the tag to the end of the polypeptide chain and then releasing the protein to be degraded. Thus, the function of the SsrA RNA and SsrA tag is to provide a method for the release of stalled ribosomes and the degradation of incomplete proteins.

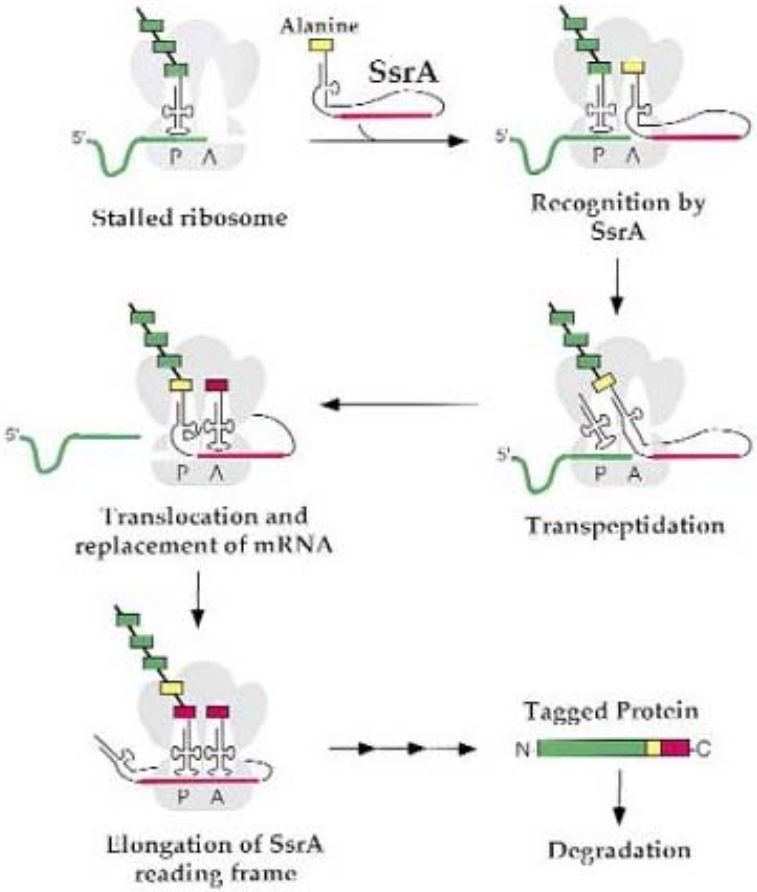

Figure 1.12. Model of SsrA tagging of proteins. An aminoacylated-SsrA RNA is recruited to a stalled ribosome which then transfers the polypeptide chain to its alanine-charged (yellow) tRNA domain. The open reading frame within SsrA (magenta) is then translated until a stop codon is reached, and the tagged protein is released. Figure adapted from [122].

1.4.2.D. Substrate Unfolding

Previous studies have shown that degradation of SsrA-tagged substrates can be increased with the overexpression of SspB [124, 125], an adaptor protein that recognizes the SsrA tag and associates with ClpXP [126]. SspB has a substrate-binding domain that contacts the N-terminal residues of the SsrA tag and a C-terminal XB peptide motif that is involved in flexible tethering to ClpX [127] (see Figure 1.13 [127]). Therefore, ClpX only contacts the C-terminal residues of the tag [128]. The SspB adaptor that recognizes

the tag then dimerizes with another SspB adaptor to deliver the tagged substrate to ClpX using the two XB peptide motifs to anchor the dimerized adaptor to ClpX.

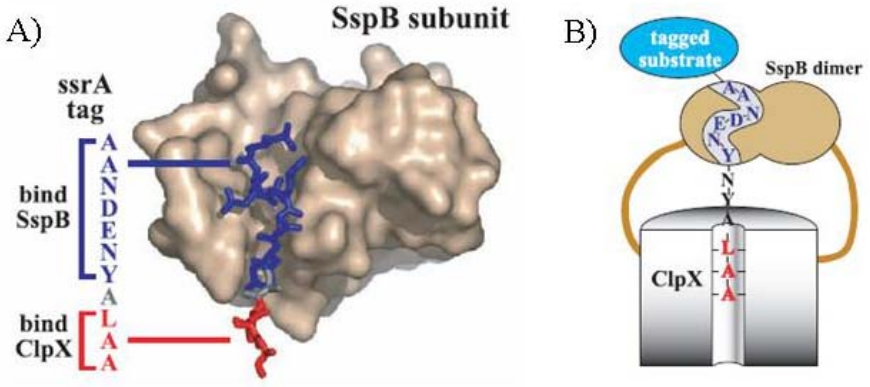

Figure 1.13. SspB binding and delivery of the SsrA tag. A) SsrA tag bound by an SspB subunit and ClpX. B) Cartoon showing the tethering of the SspB dimer to ClpX by the XB peptide motifs on flexible peptide chains. Figure adapted from [127].

In the currently proposed model, ClpA and ClpX apply mechanical force to the degradation tag causing the unfolding of the protein structure neighboring the tag [129, 130]. For ClpX, this is accomplished by the interaction between the dimerized SspB adaptor protein and ClpX itself once the adaptor is tethered by its flexible polypeptide chains (see Figure 1.14 [131]). ATP hydrolysis causes ClpX to apply a pulling force on the SsrA tag and the protein. However, this usually results in the release of the tag indicating that ATP is not used efficiently [131, 132]. The result is that either SspB attempts to feed the tag again into ClpX or disassociates completely. If the tag does not slip, the SspB dimer disengages from ClpX as more of the tag is pulled into ClpX, thus forcing SspB to release its recognized part of the tag. This release and reengagement mechanism prevents ClpX from being stalled on very stable substrates thereby making sure denatured or less stable proteins are degraded first [133].

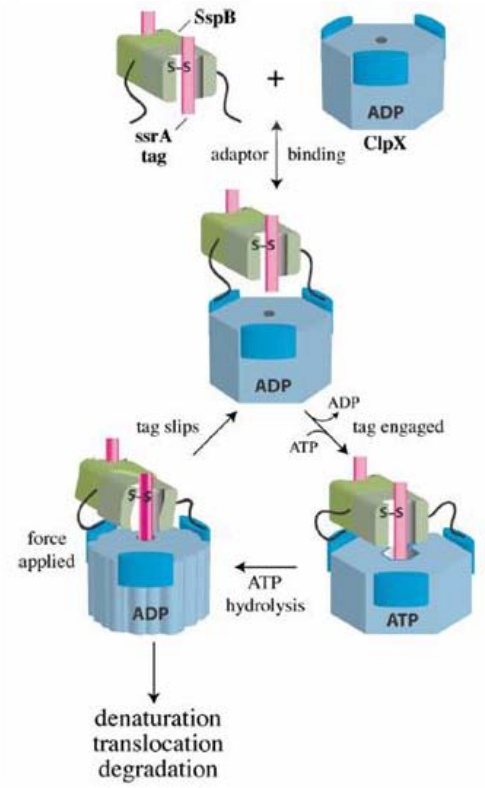

Figure 1.14. Model of the mechanistic action of ClpX. After the SspB dimer associates with ClpX, ATP hydrolysis causes ClpX to apply a force to the tag, straining its association with the dimer. This can cause slippage in which the dimer can retry or dissociate. If no slippage occurs, the tag is pulled into ClpX causing the release of the dimer. Shown is a substrate that was disulfide-linked to the dimer specifically for this research study. Figure adapted from [131].

Recent work has shown that there are three alpha-helical loops in the axial pore of ClpA that are involved in binding the SsrA tag and unfolded polypeptide chains [134]. Two loops are located at what is called the D1 ATPase domain, while the third is near the D2 ATPase domain [135]. For unfolding to occur, the tagged protein is held in the N-terminal domains of ClpA such that the tag extends axially down through the pore (see Figure 1.15 [134]). In this manner, the tag is held in place by the loops in D1 and can contact the loop in D2. The binding sites of the three loops is still unknown, but is has been proposed that they may form a shallow grove to bind the SsrA tag similar to SspB [136]. Upon ATP hydrolysis by the D2 ATPase, the D2 loop moves to a distal position

in the channel, thereby applying a mechanical force on the tag and unfolds the substrate while simultaneously translocating the polypeptide chain into the pore. The loops at D1 may provide a means to keep the unfolded chain from slipping while the D2 loop cycles back to a proximal position to bind the polypeptide chain again due to evidence that ClpA recognizes unfolded substrates without a tag element [137]. As seen with ClpX, multiple rounds of ATP hydrolysis are necessary to unfold the tagged substrate [98].

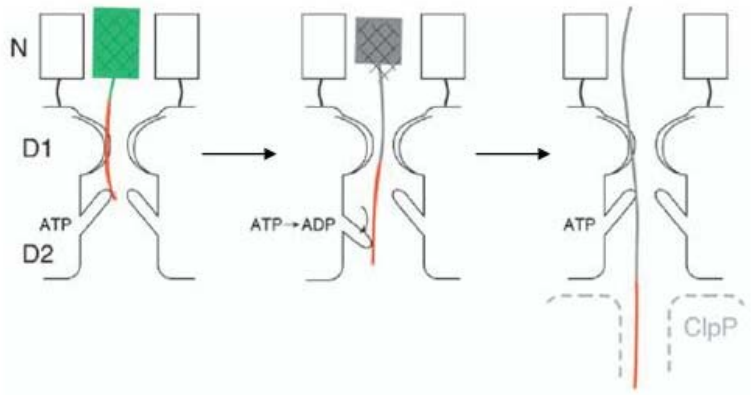

Figure 1.15. Model of the mechanistic action of ClpA. The SsrA tag (red) interacts with the three loops of ClpA, two at the D1 ATPase and one at the D2 ATPase. The tagged protein (green) is held in place by the N-terminal domains of ClpA. ATP hydrolysis by the D2 ATPase moves the D2 loop distally, applying a mechanical force that translocates the polypeptide chain and unfolds the substrate. After the D2 loop cycles, further rounds of binding, translocation and unfolding can occur with the hydrolysis of ATP. Figure adapted from [134].

1.4.2.E. Substrate Proteolysis

The proteolytic complex is typically comprised of one subunit of ClpA or ClpX per subunit of ClpP. Since the optimal ratio of ClpA or ClpX to ClpP is two to one [90], two units of either chaperone may bind to both ends of ClpP [138]. However, once a substrate begins translocation into ClpP, the ClpA or ClpX subunits bound to the either face will dissociate so as not to inhibit the translocation [139]. The affinity of ClpA for ClpP is about twice that of ClpX [113], but due to conformational changes in the IGF

loops of ClpX upon substrate binding, the affinity of ClpX to ClpP increases significantly [140].

In either case, after the protein is unfolded by ClpX or ClpA, the small linearized piece of protein is then axially fed into ClpP, and ClpA or ClpX unfolds more of the protein, repeating the cycle [139] to degrade the protein into peptides of seven to ten amino acids without any sequence specificity [104]. This tight distribution of degraded pieces occurs since translocation by the chaperone alternates with degradation by ClpP through some yet to be determined means of communication between the chaperone and ClpP [141].

ATP requirements for protein unfolding and degradation depend typically on how stable the substrate's terminus is adjacent to the tag [132]. Thus, proteins with less structure are easier to degrade. For example, research on another protease has shown that proteins with intermolecular disulfide bonds are degraded an order of magnitude slower, but typically, different variants of the same protein only vary over a two-fold rate range [142].

Three models have arisen from previous work to explain how the ClpA and ClpX ATPase motors function [143, 144]. In some studies dealing with ATPase motors similar to those of the Clp system, several motor structures have been isolated with ATP bound to all of the subunits [145, 146]. This indicates that each subunit of the Clp ATPases hydrolyzes ATP simultaneously to function as a concerted mechanism (see Figure 1.16 [147]). However, other studies show structures with ATP bound to only some of the subunits [148]. From these studies a sequential mechanism has been proposed in which the subunits hydrolyze ATP one at a time in a predetermined, sequential order around the hexamer ring. Similarly, the final model proposes that each subunit hydrolyzes ATP individually but not in any special order [144]. Developers of the model indicate that

since each part of a substrate is conformationally different and therefore can contact any subunit in the hexamer ring, certain subunits can only interact with the substrate at a given time. The subunit with the most or the best interaction may be the one that hydrolyzes ATP and further unfolds and translocates the substrate. Sauer and coworkers constructed a ClpX hexamer with only one functioning subunit which still was able to unfold and translocate proteins and used many combinations of active and inactive subunits in the hexamer to show that the rate of ATP hydrolysis was proportional to the number of active subunits [144]. Therefore, the final model seems to be the most likely model on how the Clp ATPase motors function and utilize ATP.

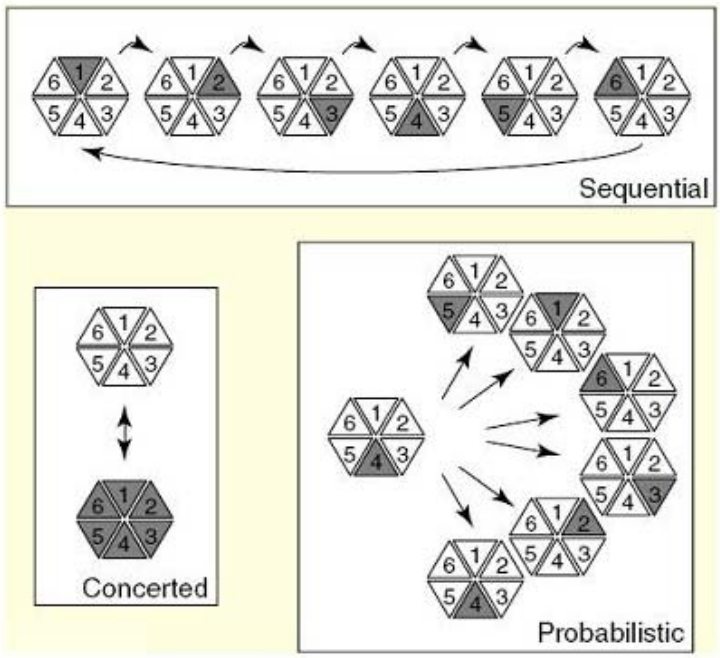

Figure 1.16. Proposed models of Clp ATPase motor function. A typical hexamer is shown and the subunit hydrolyzing ATP is shaded. The sequential model shows that one subunit hydrolyzes ATP in order around the ring. The concerted model proposes that all subunits hydrolyze simultaneously. The probabilistic model indicates that one subunit functions individually in no specified order. Figure adapted from [147].

1.5 STATE OF THE ART

1.5.1 Biotechnology and the Tat Pathway

As described above, production of recombinant proteins is typically done in bacteria, i.e. *E. coli*, since they can be grown quickly, inexpensively and can be used to express a variety of proteins. Expression of proteins in the cytoplasm allows the protein to take advantage of the molecular chaperones, cofactors, and foldases present in the cytoplasm, if necessary, to ensure proper folding. However, export of folded proteins into the periplasm via the Tat pathway could improve yields, reduce further downstream processing that would be required to remove cytoplasmic components and could be exploited for protein display purposes in the engineering of proteins with new or improved function.

However, two obstacles must be overcome in order to utilize the Tat pathway for protein production. First, some therapeutic proteins contain disulfide bonds which are necessary for proper folding and can only form in an oxidizing environment. This problem has been solved by the use of a strain of *E. coli* that has been engineered to have an oxidizing cytoplasm. Previous work has shown that disulfide-containing proteins can be exported by the Tat pathway in this oxidizing strain [55]. A recent study showed that export of a variable form of tissue plasminogen activator which contains nine disulfide bonds through the Tat pathway in the oxidizing strain yields higher amounts of active enzyme in the periplasm compared to export through the Sec pathway [149].

The second and more important barrier to using the Tat pathway for protein production is the fact that efficiency [150] and protein yields through the Tat pathway are much lower than those through the Sec pathway, which, for antibody fragments, typically exceed 1 g/L [151, 152]. This may be due to the fact that the export machinery can be

saturated by overexpressed target proteins which occurs for even native Tat export proteins [153, 154]. The saturation of the export machinery can be partially relieved by co-expression of other proteins [154, 155], however, it may also be necessary to engineer the protein to fold or interact better with the Tat export machinery or overexpress some or all of the Tat proteins in order to increase periplasmic yields.

Besides the benefit of export to the periplasm, the Tat pathway also offers a quality control advantage in that it exports only correctly folded proteins [55]. This is especially important for antibody fragment production since the Tat pathway will ensure that only correctly folded fragments in which disulfide bond formation has occurred is transported to the periplasm.

1.5.2 Engineering Protein Stability

Due to the increased use of enzymes in industrial processes, many of these biocatalysts have to be engineered with improved properties [156]. Many factors affect the stability of these enzymes such as the use of surfactants and solvents, pH, temperature and oxidative stress [157]. Many enzymes unfold and become inactivated at high temperatures, and therefore, thermostability is widely studied since the majority of industrial processes operate at high temperatures in order to maintain solubility, achieve high conversion rates, etc [1].

Several methods have been used to engineer enzymes with higher temperature stabilities. For example, the introduction of internal disulfide bonds dramatically stabilizes the protein against unfolding [158, 159]. In other studies, stability was increased by helix optimization whereby residues were changed to increase the propensity of alpha helix formation [160] and allow better interaction between neighboring alpha helices [161]. Localized residues that form salt bridges were also shown to increase the stability [162]. Finally, aromatic residues that extend from the

protein can interact if in close proximity. This "aromatic stacking" has been shown to also increase a protein's stability [163].

As indicated, there are several methods to engineer a protein for increased stability, however, none deals with increasing a protein's stability to mechanical unfolding. A protein used in flow conditions, such as in arteries may encounter a shear stress that would strain a protein's conformation. Studies have been conducted to determine a protein's stability to mechanical unfolding and will be discussed in Chapter 4 as they specifically apply to research aimed at capitalizing on the mechanism of degradation by the *E. coli* Clp protease system.

1.6 RESEARCH OBJECTIVES

Unless only a few specific changes are made to a protein to increase stability or export, a high-throughput screening method must be used to explore many combinations of multiple and/or simultaneous changes. High-throughput screening via cell display allows the linkage of genotype to phenotype required by directed evolution and the rapid measurement and screening of phenotypes from mutant libraries.

The recently discovered Tat system allows the export of fully-folded proteins across the inner cell membrane. Previously, secreted proteins were required to fold in the periplasmic space after export thus limiting the number of recombinant proteins that could be produced in *E. coli* to those that only fold correctly in the periplasm. The Tat system now allows protein engineers to take advantage of the available protein folding machinery in the cytoplasm, increasing the number of recombinant proteins that can be produced in bacteria. However, since efficiency and yield is generally low, we sought to examine the effect of mutations within the mature protein on the efficiency of export via the Tat apparatus in order to increase protein yield in the periplasm. Directed evolution and FACS was used to increase export of 26-10 scFv in an oxidizing strain of *E. coli*. The 26-10 antibody binds digoxin, a cardiac glycoside that regulates electrical impulses in the heart and the strength of contractions. One isolated clone showed increased export of 26-10 to the periplasm. *In vitro* analysis of the thermodynamic stability and folding kinetics indicated that the mutant scFv exhibits faster folding kinetics, and this is likely to be responsible for the improved export to the periplasmic space. This work is described in Chapter 2.

In addition, in order further increase the export of the 26-10 scFv through the Tat pathway, Chapter 3 describes work dealing with the co-expression of certain chaperone and Tat translocon proteins that were previously shown to relieve some saturation of the

Tat pathway. The effect of media, pH and temperature and the use of different oxidizing strains of *E. coli* on cell fluorescence and growth were also investigated.

Finally, I have used flow cytometry in an attempt to develop a genetic system for the better understanding of the Clp system and for studying how amino acid substitutions might affect the mechanical stability of protein substrates. It was proposed that since the Clp system degrades proteins, finding ways to inhibit the system would increase the yields of recombinant proteins produced in *E. coli*. In addition, since the system applies a mechanical force to unfold proteins, the Clp system could be used as an assay to engineer more stable proteins against mechanical denaturation. In order to accomplish this, GFP tagged with the SsrA peptide was used as a fluorescent reporter to develop an appropriate FACS assay as described in Chapter 4.

1.7 REFERENCES

1. Eijsink, V.G., et al., *Rational engineering of enzyme stability.* J Biotechnol, 2004. **113**(1-3): p. 105-20.
2. Presta, L.G., *Engineering of therapeutic antibodies to minimize immunogenicity and optimize function.* Adv Drug Deliv Rev, 2006.
3. Drews, J., *Drug discovery: a historical perspective.* Science, 2000. **287**(5460): p. 1960-4.
4. Vitetta, E.S. and V.F. Ghetie, *Immunology. Considering therapeutic antibodies.* Science, 2006. **313**(5785): p. 308-9.
5. Drews, J., *Stategic trends in the drug industry.* Drug Discov Today, 2003. **8**(9): p. 411-20.
6. Sharkey, R.M. and D.M. Goldenberg, *Targeted therapy of cancer: new prospects for antibodies and immunoconjugates.* CA Cancer J Clin, 2006. **56**(4): p. 226-43.
7. Carter, P.J., *Potent antibody therapeutics by design.* Nat Rev Immunol, 2006. **6**(5): p. 343-57.
8. Reichert, J.M., et al., *Monoclonal antibody successes in the clinic.* Nat Biotechnol, 2005. **23**(9): p. 1073-8.
9. Orlandi, R., et al., *Cloning immunoglobulin variable domains for expression by the polymerase chain reaction.* Proc Natl Acad Sci U S A, 1989. **86**(10): p. 3833-7.
10. Raag, R. and M. Whitlow, *Single-chain Fvs.* Faseb J, 1995. **9**(1): p. 73-80.
11. Glockshuber, R., et al., *A comparison of strategies to stabilize immunoglobulin Fv-fragments.* Biochemistry, 1990. **29**(6): p. 1362-7.
12. Arbabi-Ghahroudi, M., J. Tanha, and R. MacKenzie, *Prokaryotic expression of antibodies.* Cancer Metastasis Rev, 2005. **24**(4): p. 501-19.
13. Fromant, M., S. Blanquet, and P. Plateau, *Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction.* Anal Biochem, 1995. **224**(1): p. 347-53.
14. Cadwell, R.C. and G.F. Joyce, *Mutagenic PCR.* PCR Methods Appl, 1994. **3**(6): p. S136-40.
15. Arnold, F. and G. Georgiou, *Directed enzyme evolution: Screening and selection methods*. Methods in molecular biology. Vol. 230. 2003, Totowa, NJ.
16. Tyo, K.E., H. Zhou, and G.N. Stephanopoulos, *High-throughput screen for poly-3-hydroxybutyrate in Escherichia coli and Synechocystis sp. strain PCC6803.* Appl Environ Microbiol, 2006. **72**(5): p. 3412-7.
17. Aharoni, A., A.D. Griffiths, and D.S. Tawfik, *High-throughput screens and selections of enzyme-encoding genes.* Curr Opin Chem Biol, 2005. **9**(2): p. 210-6.
18. Olsen, M., B. Iverson, and G. Georgiou, *High-throughput screening of enzyme libraries.* Curr Opin Biotechnol, 2000. **11**(4): p. 331-7.
19. Olsen, M.J., et al., *High-throughput FACS method for directed evolution of substrate specificity.* Methods Mol Biol, 2003. **230**: p. 329-42.
20. Alberts, B., et al., *Molecular Biology of the Cell*. 2002.

21. Wulfing, C. and A. Pluckthun, *Protein folding in the periplasm of Escherichia coli.* Mol Microbiol, 1994. **12**(5): p. 685-92.
22. Thomas, J.G., A. Ayling, and F. Baneyx, *Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold.* Appl Biochem Biotechnol, 1997. **66**(3): p. 197-238.
23. Feldhaus, M.J., et al., *Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library.* Nat Biotechnol, 2003. **21**(2): p. 163-70.
24. Georgiou, G., *Analysis of large libraries of protein mutants using flow cytometry.* Adv Protein Chem, 2000. **55**: p. 293-315.
25. Chen, G., et al., *Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS).* Nat Biotechnol, 2001. **19**(6): p. 537-42.
26. Mattanovich, D. and N. Borth, *Applications of cell sorting in biotechnology.* Microb Cell Fact, 2006. **5**: p. 12.
27. Murphy, K. and A. Boyet. *FACS Lab*. 2004 [cited; Available from: http://www.pathology.wustl.edu/html/facilitiesfacs.html.
28. Perez, O.D. and G.P. Nolan, *Simultaneous measurement of multiple active kinase states using polychromatic flow cytometry.* Nat Biotechnol, 2002. **20**(2): p. 155-62.
29. Herzenberg, L.A., et al., *The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford.* Clin Chem, 2002. **48**(10): p. 1819-27.
30. Biosciences, B. *BD FACSCalibur Flow Cytometer*. 2004 [cited; Available from: www.bdbiosciences.com.
31. Mastrobattista, E., et al., *High-throughput screening of enzyme libraries: in vitro evolution of a beta-galactosidase by fluorescence-activated sorting of double emulsions.* Chem Biol, 2005. **12**(12): p. 1291-300.
32. Leary, J.F., *Strategies for rare cell detection and isolation.* Methods Cell Biol, 1994. **42 Pt B**: p. 331-58.
33. Daugherty, P.S., et al., *Antibody affinity maturation using bacterial surface display.* Protein Eng, 1998. **11**(9): p. 825-32.
34. Daugherty, P.S., et al., *Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2029-34.
35. Boder, E.T., K.S. Midelfort, and K.D. Wittrup, *Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity.* Proc Natl Acad Sci U S A, 2000. **97**(20): p. 10701-5.
36. Walsh, G., *Biopharmaceuticals: recent approvals and likely directions.* Trends Biotechnol, 2005. **23**(11): p. 553-8.
37. Manheim, B.S., Jr., P. Granahan, and K.J. Dow, *'Follow-on biologics': ensuring continued innovation in the biotechnology industry.* Health Aff (Millwood), 2006. **25**(2): p. 394-404.
38. Neidhardt, F.C., *Escherichia coli and Salmonella*. *Escherichia coli* and *Salmonella*, ed. F.C. Neidhardt. Vol. Volume 1. 1996, Washington: ASM Press.

39. McGettrick, A.F. and D.M. Worrall, *Extraction of Recombinant Protein From Bacteria*. Methods in Molecular Biology. Vol. 244. 2003.
40. Georgiou, G. and L. Segatori, *Preparative expression of secreted proteins in bacteria: status report and future prospects.* Curr Opin Biotechnol, 2005. **16**(5): p. 538-45.
41. Swartz, J.R., *Advances in Escherichia coli production of therapeutic proteins.* Curr Opin Biotechnol, 2001. **12**(2): p. 195-201.
42. Villaverde, A. and M.M. Carrio, *Protein aggregation in recombinant bacteria: biological role of inclusion bodies.* Biotechnol Lett, 2003. **25**(17): p. 1385-95.
43. Baneyx, F. and M. Mujacic, *Recombinant protein folding and misfolding in Escherichia coli.* Nat Biotechnol, 2004. **22**(11): p. 1399-408.
44. Mori, H. and K. Ito, *The Sec protein-translocation pathway.* Trends Microbiol, 2001. **9**(10): p. 494-500.
45. Veenendaal, A.K., C. van der Does, and A.J. Driessen, *The protein-conducting channel SecYEG.* Biochim Biophys Acta, 2004. **1694**(1-3): p. 81-95.
46. Kim, J. and D.A. Kendall, *Sec-dependent protein export and the involvement of the molecular chaperone SecB.* Cell Stress Chaperones, 2000. **5**(4): p. 267-75.
47. Pohlschroder, M., M.I. Gimenez, and K.F. Jarrell, *Protein transport in Archaea: Sec and twin arginine translocation pathways.* Curr Opin Microbiol, 2005. **8**(6): p. 713-9.
48. Bessonneau, P., et al., *The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure.* Embo J, 2002. **21**(5): p. 995-1003.
49. Uchida, K., H. Mori, and S. Mizushima, *Stepwise movement of preproteins in the process of translocation across the cytoplasmic membrane of Escherichia coli.* J Biol Chem, 1995. **270**(52): p. 30862-8.
50. Barrett, C.M., et al., *Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in Escherichia coli.* Biochem Biophys Res Commun, 2003. **304**(2): p. 279-84.
51. Hayhurst, A. and G. Georgiou, *High-throughput antibody isolation.* Curr Opin Chem Biol, 2001. **5**(6): p. 683-9.
52. Thomas, J.G. and F. Baneyx, *Protein folding in the cytoplasm of Escherichia coli: requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines.* Mol Microbiol, 1996. **21**(6): p. 1185-96.
53. Schatz, G. and B. Dobberstein, *Common principles of protein translocation across membranes.* Science, 1996. **271**(5255): p. 1519-26.
54. Robinson, C., C. Woolhead, and W. Edwards, *Transport of proteins into and across the thylakoid membrane.* J Exp Bot, 2000. **51 Spec No**: p. 369-74.
55. DeLisa, M.P., D. Tullman, and G. Georgiou, *Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6115-20.
56. Wickner, W. and R. Schekman, *Protein translocation across biological membranes.* Science, 2005. **310**(5753): p. 1452-6.
57. Rodrigue, A., et al., *Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway.* J Biol Chem, 1999. **274**(19): p. 13223-8.

58. Feilmeier, B.J., et al., *Green fluorescent protein functions as a reporter for protein localization in Escherichia coli.* J Bacteriol, 2000. **182**(14): p. 4068-76.
59. Tinker, J.K., J.L. Erbe, and R.K. Holmes, *Characterization of fluorescent chimeras of cholera toxin and Escherichia coli heat-labile enterotoxins produced by use of the twin arginine translocation system.* Infect Immun, 2005. **73**(6): p. 3627-35.
60. Stanley, N.R., et al., *Escherichia coli strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope.* J Bacteriol, 2001. **183**(1): p. 139-44.
61. Hutcheon, G.W. and A. Bolhuis, *The archaeal twin-arginine translocation pathway.* Biochem Soc Trans, 2003. **31**(Pt 3): p. 686-9.
62. Rose, R.W., et al., *Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway.* Mol Microbiol, 2002. **45**(4): p. 943-50.
63. Palmer, T., F. Sargent, and B.C. Berks, *Export of complex cofactor-containing proteins by the bacterial Tat pathway.* Trends Microbiol, 2005. **13**(4): p. 175-80.
64. Stanley, N.R., T. Palmer, and B.C. Berks, *The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli.* J Biol Chem, 2000. **275**(16): p. 11591-6.
65. Berks, B.C., F. Sargent, and T. Palmer, *The Tat protein export pathway.* Mol Microbiol, 2000. **35**(2): p. 260-74.
66. Cristobal, S., et al., *Competition between Sec- and TAT-dependent protein translocation in Escherichia coli.* Embo J, 1999. **18**(11): p. 2982-90.
67. Wexler, M., et al., *TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export.* J Biol Chem, 2000. **275**(22): p. 16717-22.
68. Blaudeck, N., et al., *Isolation and characterization of bifunctional Escherichia coli TatA mutant proteins that allow efficient tat-dependent protein translocation in the absence of TatB.* J Biol Chem, 2005. **280**(5): p. 3426-32.
69. De Leeuw, E., et al., *Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway.* FEBS Lett, 2001. **506**(2): p. 143-8.
70. Lee, P.A., D. Tullman-Ercek, and G. Georgiou, *The Bacterial Twin-Arginine Translocation Pathway.* Annu Rev Microbiol, 2006.
71. Lee, P.A., et al., *Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation.* J Bacteriol, 2002. **184**(21): p. 5871-9.
72. Alami, M., et al., *Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli.* Mol Cell, 2003. **12**(4): p. 937-46.
73. Barrett, C.M. and C. Robinson, *Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin-arginine translocase.* Febs J, 2005. **272**(9): p. 2261-75.
74. Palmer, T. and B.C. Berks, *Moving folded proteins across the bacterial cell membrane.* Microbiology, 2003. **149**(Pt 3): p. 547-56.

75. Ki, J.J., et al., *A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis.* J Mol Biol, 2004. **341**(4): p. 901-9.
76. Cline, K. and H. Mori, *Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport.* J Cell Biol, 2001. **154**(4): p. 719-29.
77. Mori, H. and K. Cline, *A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase.* J Cell Biol, 2002. **157**(2): p. 205-10.
78. Jack, R.L., et al., *Constitutive expression of Escherichia coli tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth.* J Bacteriol, 2001. **183**(5): p. 1801-4.
79. Berks, B.C., T. Palmer, and F. Sargent, *The Tat protein translocation pathway and its role in microbial physiology.* Adv Microb Physiol, 2003. **47**: p. 187-254.
80. Gohlke, U., et al., *The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10482-6.
81. Oates, J., et al., *Consensus structural features of purified bacterial TatABC complexes.* J Mol Biol, 2003. **330**(2): p. 277-86.
82. de Leeuw, E., et al., *Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes.* J Mol Biol, 2002. **322**(5): p. 1135-46.
83. Musser, S.M. and S.M. Theg, *Proton transfer limits protein translocation rate by the thylakoid DeltapH/Tat machinery.* Biochemistry, 2000. **39**(28): p. 8228-33.
84. Lodish, H.B., Arnold; Zipursky, S. Lawrence; Matsudaira, Paul; Baltimore, David; Darnell, James E., *The Proton-Motive Force in Mitochondria Is Due Largely to a Voltage Gradient across the Inner Membrane*. Molecular Cell Biology. 2000.
85. Alder, N.N. and S.M. Theg, *Energy use by biological protein transport pathways.* Trends Biochem Sci, 2003. **28**(8): p. 442-51.
86. Jenal, U. and T. Fuchs, *An essential protease involved in bacterial cell-cycle control.* Embo J, 1998. **17**(19): p. 5658-69.
87. Jenal, U. and R. Hengge-Aronis, *Regulation by proteolysis in bacterial cells.* Curr Opin Microbiol, 2003. **6**(2): p. 163-72.
88. Straus, D.B., W.A. Walter, and C.A. Gross, *Escherichia coli heat shock gene mutants are defective in proteolysis.* Genes Dev, 1988. **2**(12B): p. 1851-8.
89. Porankiewicz, J., J. Wang, and A.K. Clarke, *New insights into the ATP-dependent Clp protease: Escherichia coli and beyond.* Mol Microbiol, 1999. **32**(3): p. 449-58.
90. Maurizi, M.R., et al., *Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP.* Biochemistry, 1998. **37**(21): p. 7778-86.
91. Gottesman, S., S. Wickner, and M.R. Maurizi, *Protein quality control: triage by chaperones and proteases.* Genes Dev, 1997. **11**(7): p. 815-23.
92. Gottesman, S., *Proteases and their targets in Escherichia coli.* Annu Rev Genet, 1996. **30**: p. 465-506.

93. Gottesman, S. and M.R. Maurizi, *Regulation by proteolysis: energy-dependent proteases and their targets.* Microbiol Rev, 1992. **56**(4): p. 592-621.
94. Gottesman, S., *Proteolysis in bacterial regulatory circuits.* Annu Rev Cell Dev Biol, 2003. **19**: p. 565-87.
95. Hwang, B.J., et al., *Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La.* Proc Natl Acad Sci U S A, 1987. **84**(16): p. 5550-4.
96. Maurizi, M.R., et al., *Clp P represents a unique family of serine proteases.* J Biol Chem, 1990. **265**(21): p. 12546-52.
97. Gottesman, S., et al., *ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities.* J Biol Chem, 1993. **268**(30): p. 22618-26.
98. Lee, C., et al., *ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal.* Mol Cell, 2001. **7**(3): p. 627-37.
99. Thompson, M.W., S.K. Singh, and M.R. Maurizi, *Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis.* J Biol Chem, 1994. **269**(27): p. 18209-15.
100. Singh, S.K., et al., *Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8898-903.
101. Szyk, A. and M.R. Maurizi, *Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site.* J Struct Biol, 2006.
102. Wang, J., J.A. Hartling, and J.M. Flanagan, *The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis.* Cell, 1997. **91**(4): p. 447-56.
103. Burton, B.M. and T.A. Baker, *Remodeling protein complexes: insights from the AAA+ unfoldase ClpX and Mu transposase.* Protein Sci, 2005. **14**(8): p. 1945-54.
104. Thompson, M.W. and M.R. Maurizi, *Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates.* J Biol Chem, 1994. **269**(27): p. 18201-8.
105. Gribun, A., et al., *The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation.* J Biol Chem, 2005. **280**(16): p. 16185-96.
106. Banecki, B., et al., *Structure-function analysis of the zinc-binding region of the Clpx molecular chaperone.* J Biol Chem, 2001. **276**(22): p. 18843-8.
107. Smith, C.K., T.A. Baker, and R.T. Sauer, *Lon and Clp family proteases and chaperones share homologous substrate-recognition domains.* Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6678-82.
108. Bochtler, M., et al., *The structures of HsIU and the ATP-dependent protease HsIU-HsIV.* Nature, 2000. **403**(6771): p. 800-5.
109. Singh, S.K., et al., *Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis.* J Biol Chem, 2001. **276**(31): p. 29420-9.

110. Kim, D.Y. and K.K. Kim, *Crystal structure of ClpX molecular chaperone from Helicobacter pylori.* J Biol Chem, 2003. **278**(50): p. 50664-70.
111. Kim, Y.I., et al., *Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase.* Nat Struct Biol, 2001. **8**(3): p. 230-3.
112. Katayama, Y., et al., *The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component.* J Biol Chem, 1988. **263**(29): p. 15226-36.
113. Grimaud, R., et al., *Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP.* J Biol Chem, 1998. **273**(20): p. 12476-81.
114. Pak, M., et al., *Concurrent chaperone and protease activities of ClpAP and the requirement for the N-terminal ClpA ATP binding site for chaperone activity.* J Biol Chem, 1999. **274**(27): p. 19316-22.
115. Farrell, C.M., A.D. Grossman, and R.T. Sauer, *Cytoplasmic degradation of ssrA-tagged proteins.* Mol Microbiol, 2005. **57**(6): p. 1750-61.
116. Flynn, J.M., et al., *Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals.* Mol Cell, 2003. **11**(3): p. 671-83.
117. Hoskins, J.R., S.Y. Kim, and S. Wickner, *Substrate recognition by the ClpA chaperone component of ClpAP protease.* J Biol Chem, 2000. **275**(45): p. 35361-7.
118. Hoskins, J.R., et al., *ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11037-42.
119. Hinnerwisch, J., et al., *Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease.* J Biol Chem, 2005. **280**(49): p. 40838-44.
120. Sharma, S., J.R. Hoskins, and S. Wickner, *Binding and degradation of heterodimeric substrates by ClpAP and ClpXP.* J Biol Chem, 2005. **280**(7): p. 5449-55.
121. Keiler, K.C., P.R. Waller, and R.T. Sauer, *Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA.* Science, 1996. **271**(5251): p. 990-3.
122. Karzai, A.W., E.D. Roche, and R.T. Sauer, *The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue.* Nat Struct Biol, 2000. **7**(6): p. 449-55.
123. Roche, E.D. and R.T. Sauer, *SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity.* Embo J, 1999. **18**(16): p. 4579-89.
124. Levchenko, I., et al., *A specificity-enhancing factor for the ClpXP degradation machine.* Science, 2000. **289**(5488): p. 2354-6.
125. McGinness, K.E., T.A. Baker, and R.T. Sauer, *Engineering controllable protein degradation.* Mol Cell, 2006. **22**(5): p. 701-7.
126. Ades, S.E., *Proteolysis: Adaptor, adaptor, catch me a catch.* Curr Biol, 2004. **14**(21): p. R924-6.

127. Hersch, G.L., T.A. Baker, and R.T. Sauer, *SspB delivery of substrates for ClpXP proteolysis probed by the design of improved degradation tags.* Proc Natl Acad Sci U S A, 2004. **101**(33): p. 12136-41.
128. Gottesman, S., et al., *The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system.* Genes Dev, 1998. **12**(9): p. 1338-47.
129. Kenniston, J.A., et al., *Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation.* J Struct Biol, 2004. **146**(1-2): p. 130-40.
130. Burton, R.E., et al., *Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine.* Embo J, 2001. **20**(12): p. 3092-100.
131. Bolon, D.N., et al., *Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease.* Mol Cell, 2004. **16**(3): p. 343-50.
132. Kenniston, J.A., et al., *Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine.* Cell, 2003. **114**(4): p. 511-20.
133. Kenniston, J.A., T.A. Baker, and R.T. Sauer, *Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing.* Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1390-5.
134. Hinnerwisch, J., et al., *Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation.* Cell, 2005. **121**(7): p. 1029-41.
135. Xia, D., et al., *Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone.* J Struct Biol, 2004. **146**(1-2): p. 166-79.
136. Levchenko, I., et al., *Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag.* Mol Cell, 2003. **12**(2): p. 365-72.
137. Hoskins, J.R., et al., *Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8892-7.
138. Ortega, J., et al., *ClpA and ClpX ATPases bind simultaneously to opposite ends of ClpP peptidase to form active hybrid complexes.* J Struct Biol, 2004. **146**(1-2): p. 217-26.
139. Ortega, J., et al., *Alternating translocation of protein substrates from both ends of ClpXP protease.* Embo J, 2002. **21**(18): p. 4938-49.
140. Joshi, S.A., et al., *Communication between ClpX and ClpP during substrate processing and degradation.* Nat Struct Mol Biol, 2004. **11**(5): p. 404-11.
141. Choi, K.H. and S. Licht, *Control of peptide product sizes by the energy-dependent protease ClpAP.* Biochemistry, 2005. **44**(42): p. 13921-31.
142. Kwon, A.R., C.B. Trame, and D.B. McKay, *Kinetics of protein substrate degradation by HslUV.* J Struct Biol, 2004. **146**(1-2): p. 141-7.
143. Ogura, T. and A.J. Wilkinson, *AAA+ superfamily ATPases: common structure--diverse function.* Genes Cells, 2001. **6**(7): p. 575-97.
144. Martin, A., T.A. Baker, and R.T. Sauer, *Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines.* Nature, 2005. **437**(7062): p. 1115-20.

145. Gai, D., et al., *Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen.* Cell, 2004. **119**(1): p. 47-60.
146. Wang, J., et al., *Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU.* Structure, 2001. **9**(11): p. 1107-16.
147. Ades, S.E., *AAA+ molecular machines: firing on all cylinders.* Curr Biol, 2006. **16**(2): p. R46-8.
148. Singleton, M.R., et al., *Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides.* Cell, 2000. **101**(6): p. 589-600.
149. Kim, J.Y., et al., *Twin-arginine translocation of active human tissue plasminogen activator in Escherichia coli.* Appl Environ Microbiol, 2005. **71**(12): p. 8451-9.
150. Thomas, J.D., et al., *Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli.* Mol Microbiol, 2001. **39**(1): p. 47-53.
151. Barth, S., et al., *Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions.* Appl Environ Microbiol, 2000. **66**(4): p. 1572-9.
152. Cupit, P.M., et al., *Cloning and expression of single chain antibody fragments in Escherichia coli and Pichia pastoris.* Lett Appl Microbiol, 1999. **29**(5): p. 273-7.
153. Yahr, T.L. and W.T. Wickner, *Functional reconstitution of bacterial Tat translocation in vitro.* Embo J, 2001. **20**(10): p. 2472-9.
154. DeLisa, M.P., et al., *Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway.* J Bacteriol, 2004. **186**(2): p. 366-73.
155. Li, S.Y., B.Y. Chang, and S.C. Lin, *Coexpression of TorD enhances the transport of GFP via the TAT pathway.* J Biotechnol, 2006. **122**(4): p. 412-21.
156. Kirk, O., T.V. Borchert, and C.C. Fuglsang, *Industrial enzyme applications.* Curr Opin Biotechnol, 2002. **13**(4): p. 345-51.
157. Eijsink, V.G., et al., *Directed evolution of enzyme stability.* Biomol Eng, 2005. **22**(1-3): p. 21-30.
158. Mansfeld, J., et al., *Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond.* J Biol Chem, 1997. **272**(17): p. 11152-6.
159. Clarke, J. and A.R. Fersht, *Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation.* Biochemistry, 1993. **32**(16): p. 4322-9.
160. Blaber, M., X.J. Zhang, and B.W. Matthews, *Structural basis of amino acid alpha helix propensity.* Science, 1993. **260**(5114): p. 1637-40.
161. Marshall, S.A., C.S. Morgan, and S.L. Mayo, *Electrostatics significantly affect the stability of designed homeodomain variants.* J Mol Biol, 2002. **316**(1): p. 189-99.
162. Makhatadze, G.I., et al., *Contribution of surface salt bridges to protein stability: guidelines for protein engineering.* J Mol Biol, 2003. **327**(5): p. 1135-48.
163. Puchkaev, A.V., L.S. Koo, and P.R. Ortiz de Montellano, *Aromatic stacking as a determinant of the thermal stability of CYP119 from Sulfolobus solfataricus.* Arch Biochem Biophys, 2003. **409**(1): p. 52-8.

Chapter 2

Engineering Proteins to Increase Tat Export Efficiency

2.1 INTRODUCTION

As described in the first chapter, the conformation of proteins in the cytoplasm is a critical element in determining their competence for translocation across lipid bilayer membranes. In *E. coli*, proteins that are exported via the general secretory pathway must be maintained in a partially unfolded conformation that can be threaded via the narrow pore of the SecYEG translocation apparatus [1-3]. For some proteins, this is not a problem because they are synthesized co-translationally by ribosomes that dock onto the membrane via the interaction of the signal recognition particle with its receptor [4]. However, for the majority of proteins that utilize the general secretory pathway, membrane translocation occurs post-translationally. Export competence is ensured by interactions with chaperones such as SecB and by the signal peptide which serves to retard folding [5-8]. Proteins that fold too rapidly to benefit from these processes are incompetent for export and remain in the cytoplasm. Consistent with this idea, a recent genetic selection for mutants of the fast folding cytoplasmic protein thioredoxin, which is not compatible for export when fused to the signal peptide of a post-translationally exported protein, resulted in the isolation of slower folding variants [9].

In sharp contrast to export via the SecYEG pore, proteins that utilize the post-translational Tat secretion pathway have to be folded in order to be competent for export [10]. Misfolded or partially folded proteins cannot be translocated via the Tat pathway, suggesting the existence of a folding quality control mechanism that operates either prior

to, or concomitant with, translocation through the Tat pore [11, 12]. Not surprisingly polypeptides prone to aggregation cannot be exported to the periplasm via the Tat pathway. Recently, Fisher *et al.* showed that tripartite fusions comprising of: (i) the Tat specific signal peptide ssTorA; (ii) variants of the amyloid peptide Aβ42 and (iii) the reporter protein β-lactamase could be localized in the periplasm and confer resistance to β-lactam antibiotics only if the Aβ42 moiety was soluble [13]. Mutations that increased the solubility of the Aβ42 peptide domain of the tripartite fusion allowed better Tat export and therefore resulted in higher resistance to antibiotics.

The notion that unfolded proteins cannot be translocated via the Tat pathway is supported by *in vivo* and *in vitro* evidence from bacteria and from plant thylakoids [14-19]. However, the relationship between the *in vitro* folding properties of a protein and competence for Tat export has not been investigated. Processes that are dictated by the folding kinetics, such as of off-pathway reactions leading to the formation of aggregates or interactions with chaperones, are known to be important for protein translocation [20, 21].

In this study, we sought to examine the effect of mutations within the mature protein on the efficiency of export via the Tat apparatus in order to increase protein yield in the periplasm. scFv antibody fragments comprised of the V_H and the V_L immunoglobulin domains linked by a $(Gly_4Ser)_3$ are widely used for biotechnology applications, and their folding characteristics have been studied in detail [22-24]. The 26-10 scFv antibody fragment binds to digoxin and to other cardiac glycosides that regulate electrical impulses in the heart and the strength of contractions with nanomolar affinity [25, 26]. By using a fluorescently labeled derivative of digoxin and PECS (described in Chapter 1), high-throughput sorting was carried out on mutant libraries of the 26-10 scFv to isolate clones with increased fluorescence. One isolated clone with

approximately twice the fluorescence of wild type (WT) 26-10 scFv showed increased export of 26-10 to the periplasm. *In vitro* analysis of the thermodynamic stability and folding kinetics indicated that the mutant scFv exhibits faster folding kinetics. The faster folding rate is likely to be responsible for the improved export to the periplasmic space.

Special thanks to Thomas Van Blarcom for help with sorting by MoFlow, Christian Cobaugh and Mark Gebhard for help with BIACore experiments and Clint Leysath for help with protein purification.

2.2 MATERIALS AND METHODS

2.2.1 Strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table 2.1. *E. coli* XL1-Blue cells were used for plasmid propagation.

To construct the p99PelB-2610 pTrc99A plasmid, primers PelB.f, which contains the whole *pelB* signal peptide, and PelB.r were used to amplify the *26-10* gene from p99TorA-2610 by PCR. The PCR fragment was digested with *Bsph*I and *Xma*I and empty pTrc99A plasmid was digested by *Nco*I and *Xma*I. Since *Bsph*I and *Nco*I leave the same sticky ends of DNA upon cleavage, ligation of the PCR fragment and the digested pTrc99A plasmid forms the p99PelB-2610 construct.

Using primers 14B7.f and 14B7.r and p99-14B7 as template, the *14B7* gene was amplified by PCR. p99TorA-2610 and p99PelB-2610 DNA was digested with *Nhe*I and then treated with calf intestinal alkaline phosphatase (CIAP) to prevent the digested plasmid backbone from ligating to itself during the ligation step. The PCR products were also digested with *Nhe*I and then ligated into both digested plasmids separately, creating p99TorA-14B7 and p99PelB-14B7.

The pTorA-2610 pBAD18 scFv fusion was cloned using overlap extension PCR with primer sets T2610-1 to T2610-3. Primers T2610-2.f and T2610-2.r in the overlap PCR changed a *NheI* restriction site after the *torA* signal peptide to an *XbaI* restriction site and primers T2610-3.f and T2610-3.r removed an *XbaI* restriction site internal to 26-10 with a silent mutation. p99TorA-2610 was used as the template in the reactions. The fragment was then inserted into pBAD18 using the *XmaI* and *SphI* restriction sites. Primers P2610.f and P2610.r were used with p99PelB-2610 as template to clone the *pelB* signal into the pTorA-2610 construct using the *XmaI* and *XbaI* restriction sites making

pPelB-2610. In the same manner, the P2610 primers were used to clone the *pelB* signal into the pTorA-2610$_{C8}$ mutant clone creating pPelB-2610$_{C8}$.

Table 2.1. Strains, plasmids, and primers used in this study

E. coli Strain or Plasmid	Relevant Genotype or Features	Source
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q} *ZΔM15* Tn*10* (Tet^r)]	Lab collection
DHB4	MC4100 *phoR* Δ(*phoA*) *Pvu*II Δ(*malF)3* F' [*lacI*^{q}*ZYA -pro*]	Lab collection
FA113	DHB4 *trxB gor552* Tn*10*tet^r *ahpC**	J. Beckwith
FA113 ΔtatC	DHB4 *trxB gor552* Tn*10*tet^r *ahpC** *tatC::spec*	Lab collection
pBAD18	*araBAD* promoter, pBR322 *ori*, cm^r	Ref. [27]
pTrc99A	*trc* promoter, ColE1 *ori*, cm^r	Amersham Biosciences
p99-14B7	14B7 in Trc99A	Lab collection
p99TorA-2610	ssTorA-26-10 in pTrc99A	Lab collection
p99TorA-14B7	ssTorA-14B7 in pTrc99A	This study
pTorA-2610	ssTorA-26-10 in pBAD18	This study
pTorA-2610$_{C8}$	ssTorA-26-10 mutant C8 in pBAD18 isolated from library	This study
p99PelB-2610	ssPelB-26-10 in pTrc99A	This study
p99PelB-14B7	ssPelB-14B7 in pTrc99A	This study
pPelB-2610	ssPelB-26-10 in pBAD18	This study
pPelB-2610$_{C8}$	ssPelB-26-10 mutant C8 in pBAD18	This study
Primers		
PelB.f	GTCTTCATGAAATACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGC GGCCCAGCCGGCCATGGCCGCTAGCGAAGTTCAACTGCAA	
PelB.r	CCTAGCCCGGGTCAGTGGTGGTGGTGGTG	
14B7.f	ATATGCTAGCGATATTCAGATGACACAGACT	
14B7.r	ATATCTCGAGCGAGGAGACGGTGAC	
T2610-1.f	GATCCCCGGGTTAAAGAGGAGAAAGGTCATGAACAATAACGATCTCTTTCAG GCA	
T2610-1.r	TAGGCATGCTCAGTGGTGGTGGTGGTGGTGCTC	
T2610-2.f	CAGACTGTTGCAGTTGAACTTCTCTAGAAGCGTCAGTCGCCGCTTG	
T2610-2.r	TCTAGAGAAGTTCAACTGCAACAGTCTG	
T2610-3.f	ATGGTAAGTCTTTAGACTACATCGGG	
T2610-3.r	TACCATTCAGAAATCTGATGTAGCCC	
P2610.f	GATCCCCGGGTTAAAGAGGAGAAAGGTCATGAAATACCTATTGCCTACGGC AGCCGCTGGA	
P2610.r	CTAGTCTAGAGGCCATGGCCGGCTGGGCCGCGAGTAATAACAATCCAGCGG CTGCCGTA	
E2610.f	GCGACTGACGCTTCTAGA	
E2610.r	AGCCCGTTTGATCTCGAG	

2.2.2 Growth conditions, expression, PECS and purification

Unless otherwise noted, cells were grown at 37°C on LB media with 25 μg/ml chloramphenicol, 25 μg/ml kanamycin and/or 100 μg/ml ampicillian, as appropriate.

Cells were grown in TB media containing the appropriate antibiotics at 37°C to an OD_{600} of approximately 0.3, and protein synthesis was induced by adding arabinose and/or IPTG to a final concentration of 0.2% and 0.1 mM, respectively. Cultures were grown for an additional 4 hours at 30°C at which point 0.1 ml aliquots were taken for membrane permeabilization, labeling and FACS sorting per the procedure described previously [28]. For labeling, cells were mixed with 0.9 ml 5xPBS (phosphate buffered saline) containing 200 nM digoxigenin-BODIPY (4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionyl ethylenediamine) [29] (Molecular Probes). The samples were incubated at room temperature for one hour with shaking, and the cells were pelleted at 2,000 g and resuspended in 1 ml of 5xPBS before being analyzed by flow cytometry. Flow cytometry was carried out on a Becton-Dickinson FACSort flow cytometer with the following settings in log format: Forward scatter E01, Side scatter trigger 350 volts, FL-1 700 volts, FL-2 700 volts. 5xPBS was run as sheath for all scans. Fluorescence was monitored by 488 nm emission via a 530/30 band-pass filter (FL-1).

Colonies isolated following sorting were picked to individual wells of a 96-well plate (Costar brand round bottom, polystyrene, non-binding surface) containing 100 μl of TB media with the appropriate antibiotics. After overnight growth at 37°C, fresh microtiter well plates were inoculated, the cells were grown for 1.5 hours at 37°C, and induced by the addition of arabinose to a final concentration of 0.2%, respectively. Cultures were then grown at 30°C for an additional four hours. Subsequently, the cells were pelleted by centrifugation at 2,000 g for ten minutes at room temperature. The cells in each well were resuspended in 200 μl of 5xPBS containing 200 nM digoxigenin-

BODIPY and incubated for one hour at room temperature. The plates were centrifuged again at 2,000 g for ten minutes at room temperature. 200 µl of 5xPBS was added to each well to resuspend the cells before the plate was analyzed by a BioTek Synergy™ HT Microplate Reader using an excitation filter of 480 nm and an emission filter of 530 nm. To confirm fluorescence readings, 5 µl of sample from a well was added to 1 ml of 5xPBS in a sieve tube to be analyzed by flow cytometry.

For mid-scale expression and purification [28, 30], 40 ml cultures were grown overnight in TB media at 37^{o}C. These cultures were added to 460 ml fresh TB media and grown until mid-log phase and induced with arabinose to 0.2% final concentration. Incubation was continued for four hours at 30^{o}C, and the cells were harvested by centrifugation at 4,000 g for 10 min at 4^{o}C. The cell paste was resuspended in 10 ml of 10 mM Tris, 0.75 M sucrose, pH 7.5. 1 ml of lysozyme was added, and the cells were swirled on ice for 20 minutes. 20 ml of 1 mM ethylenediaminetetraacetic acid (EDTA) was added slowly, and the cells were left to swirl for another 15 minutes. 1.4 ml of 0.5 M $MgCl_2$ was added and another 15 minutes of swirling elapsed. The resulting spheroplasts were then pelleted at 10,000 g for 15 minutes at 4^{o}C.

0.5 ml of Ni-NTA agarose resin (Qiagen) was washed repeatedly in 1xIMAC buffer (0.02 M Na_2HPO_4, 500 mM NaCl, 0.1 M imidazole, 0.1% Tween-20, pH 7.5) to remove the ethanol storage solution. The supernatant was decanted after centrifugation, and 3.5 ml of 10xIMAC buffer was added along with the Ni-NTA agarose. After shaking on ice for one hour, the resin was pelleted at 3,000 g for 5 minutes at 4^{o}C. The resin was washed twice with 50 ml of 1xIMAC buffer with centrifugation in between washes. 5 ml of the last wash was kept with the resin and loaded onto a BioRad polyprep disposable column to drain. 1 ml of 1 M imidazole in 1xIMAC buffer was added to the column and left for five minutes before collecting the eluate. The collected eluate was injected into

an Äkta FPLC system (Amersham Pharmacia) using a Superdex 200 HR10/30 column (Amersham Pharmacia). Fractions (0.5 ml) corresponding to the molecular weight of the scFv (as determined by running protein standards) were collected, concentrated using Amicon Ultra Cellulose filters (Millipore), and quantified by bicinchoninic acid assay (Pierce) using BSA as a standard and SDS-PAGE with known molecular standards.

2.2.3 Library construction and sorting

Primers E2610.f and E2610.r were used with pTorA-2610 as template for error-prone PCR [31] to create a library with approximately 13 mutations in the *26-10* scFv gene corresponding to a mutation rate of 1.3% determined by sequencing six clones from the library. The PCR were inserted into the pTorA-2610 plasmid using the *XbaI* and *XhoI* restriction sites. The ligation mixture was electroporated into *E. coli* XL1-Blue cells and plated on antibiotic containing plates. Plasmid was isolated from the resulting colonies by QIAprep Spin Miniprep Kit (Qiagen). The isolated plasmids were then used to transform the library into the FA113 strain resulting in $6x10^6$ transformants.

1 ml of frozen stock of library cells was used to inoculate 100 ml of TB media supplemented with 25 μg/ml chloramphenicol. After growth for 1 hour at 37°C, expression of the 26-10 scFv was induced with arabinose to a final concentration of 0.2%. Growth was continued for four hours at 30°C. 100 μl of cells were labeled as above. The library was then sorted on a MoFlo flow cytometer (FC) equipped with an Innova 90C argon laser (Cytomation, Fort Collins, CO). Triggering on side scatter was set to 488 nm light from the argon laser. Detection of digoxigenin-BODIPY binding was monitored by 518 nm emission via a 530/40 band-pass filter (FL-1). By gating for cells showing high FL-1 signal, the library of scFvs was subjected to three rounds of sorting. Between each round of sorting, the collected events were subject to colony PCR amplification using primers E2610.f and E2610.r and reinserted into the pTorA-2610

construct using the *Xba*I and *Xho*I restriction sites. Clones isolated after round 3 selection were picked to 96-well plates, scanned by fluorescent plate reader and checked by flow cytometry.

2.2.4 Biochemical Assays

Cell fractionations were performed by the cold osmotic shock procedure, as described previously [11]. Briefly, cells were harvested by centrifugation at 2,000 g and resuspended in 1 ml fractionation buffer (30 mM Tris-HCl pH 8.0, 20% w/v sucrose, 1 mM ethylenediaminetetraacetic acid, EDTA) per 10 ml of culture. After ten minutes at 25^{o}C, the cells were centrifuged at 6,000 g for 15 minutes at 4^{o}C. The pellet was resuspended in 1 ml of ice cold 5 mM $MgSO_4$ ice cold and left on ice for 5-10 minutes. After centrifugation at 16,100 g for 15 minutes at 4^{o}C, the supernatant was kept as the periplasmic fraction. The cell pellet was resuspended in 266 µl of 1xPBS, sonicated for 30 s, and recentrifuged at 16,100 g for 15 minutes at 4^{o}C to obtain the soluble cytoplasmic fraction and remove insoluble cell debris.

Western blotting was performed as previously described in Chapter 2. The primary antibody was a monoclonal mouse anti-polyhistidine peroxidase conjugate (Sigma) diluted 1:10,000. Western blots were also probed with anti-GroEL antibody (Sigma) as a control.

Enzyme-linked immunosorbent assays (ELISA) were performed as previously described [32]. Digoxigenin-BSA (bovine serum albumin) at 2 µg/ml was used to coat the plates, and each ELISA was conducted on 100 µl of periplasmic fraction. The monoclonal mouse anti-polyhistidine peroxidase conjugate (Sigma) was used to detect protein binding at a dilution of 1:3,000. o-phenylenediamine tablets were used to develop the signal, and the reaction was stopped with H_2SO_4 before the plates were read at 490 nm on a Beckman BioMek plate reader.

2.2.5 Protein Characterization

BIACore analysis was performed as previously described [33] using a BIACore 3000 instrument. Briefly, a CM-5 carboxymethylated dextran matrix chip (BIACore) was used for the immobilization of digoxigenin-BSA (approximately 130 RU, relative units) following the standard protocol from Pharmacia Biosensor. All kinetics experiments were performed in HEPES-EP buffer containing 150 mM NaCl, 10 mM HEPES, 3.4 mM EDTA, and 0.005% polysorbate-20, pH 7.4 at 25°C with a flow rate of 100 μl/min. scFv proteins in the concentration range of 0.5 to 500 nM were added to the chip for the association rate measurements. Regeneration of the antibody-bound chip was accomplished by using 50% ethylene glycol/HEPES-EP, pH 10. BIAEvaluation software from Pharmacia Biosensor was used to calculate the association and dissociation rate values of the scFvs.

Guanidine HCl (GdnHCl)-induced denaturation equilibrium experiments were carried out in degassed and sterile-filtered BBS (50 mM sodium borate, 150 mM sodium chloride) at 20^{o}C at a final protein concentration of 15 μg/ml. 15 mM dithiothreitol (DTT) was added to keep the protein reduced when necessary. Excitation wavelength was 280 nm, and the emission spectrum of the protein was scanned from 320 nm to 370 nm using an AMINCO-Bowman Series 2 Luminescence Spectrometer. Three spectra were averaged per sample, and the fraction of unfolded protein was calculated as described previously [34] by using the maxima at 360 nm.

For renaturation kinetics, proteins were first denatured overnight at 20^{o}C in BBS containing 80 mM DTT and 5.0 M GdnHCl. The samples were then diluted to 0.5 M GdnHCl using BBS to a final protein concentration of 15 μg / ml, and the time-course for folding was measured by following the increase in fluorescence at 325 nm with excitation

at 275 nm. Data points represented the average of four separate experiments. Fitted curves were used to determine the macroscopic rate constants.

For aggregation measurements of scFvs by light scattering, proteins at a final concentration of 15 μg / ml were incubated in different concentrations of GdnHCl for 16 hours at 20^{o}C, and the light scattering signal was monitored at 350 nm with emission at 350 nm. Data show the average of three independent experiments.

2.3 RESULTS

2.3.1 26-10 can be exported by the Tat pathway

Previously, our laboratory had reported that the 26-10 scFv can be exported into the periplasmic space by fusion to the Tat specific signal peptide ssTorA from the *E. coli* trimethylamine *N*-oxide reductase [11]. Export was observed when the ssTorA-26-10 scFv gene fusion was expressed only in *E. coli* strains that have an oxidizing cytoplasm, such a FA113 (*trxB gor ahpC**). Expression in *trxB gor* mutant strains allows the formation of the two disulfide bonds in scFv that are important for the stability of the protein [11, 35, 36]. The *trxB* gene of *E. coli* encodes for thioredoxin reductase which functions to maintain the reducing nature of the cytoplasm. Similarly the *gor* gene encodes for the glutathione reductase enzyme which reduces glutathiones, in turn, required for the action of glutaredoxins. The reduced forms of thioredoxins and glutathiones maintain the reducing nature of the cytoplasm by reducing disulfide bonds in cytoplasmic proteins.

The amount of 26-10 scFv in the periplasm could be monitored by flow cytometry following partial permeabilization of the outer membrane by exposure to hypertonic buffer (5xPBS) and incubation with the fluorescent hapten digoxigenin-BODIPY. Under these conditions, the fluorescent hapten diffuses readily across the outer membrane while the much larger antibody fragment cannot escape from the periplasm (see Figure 2.1). Binding of the hapten by the scFv in the periplasm results in higher cell fluorescence. Figure 2.2 shows that the cell fluorescence of *E. coli* FA113 expressing the ssTorA-26-10 scFv fusion was approximately 3 times higher than the background cell autofluorescence in cells that do not contain plasmid. Inactivation of the Tat export pathway by deletion of the essential Tat translocon component *tatC* resulted in low cell fluorescence. Similarly,

background fluorescence was detected in the isogenic parental strain *E. coli* DHB4 where the oxidative folding of the scFv in the cytoplasm is prevented, and therefore the antibody fragment is incompetent for Tat export. Since only background fluorescence occurs in strains containing pTorA-14B7, a scFv of unrelated specificity (the 14B7 scFv recognizes the PA antigen from *Bacillus anthracis*), non-specific binding of the digoxigenin-BODIPY probe did not occur.

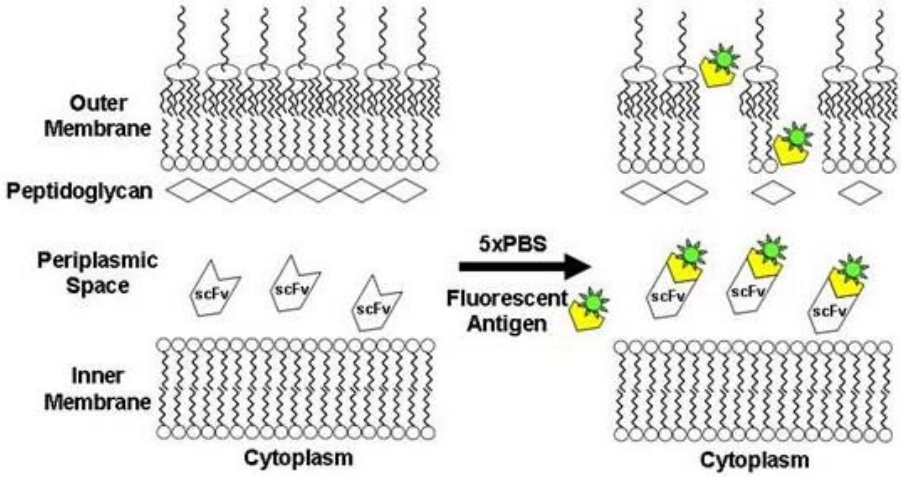

Figure 2.1. Schematic showing the detection of periplasmic scFv antibodies via binding to fluorescent hapten. 5xPBS is used to permeabilize the outer membrane to the fluorescent hapten.

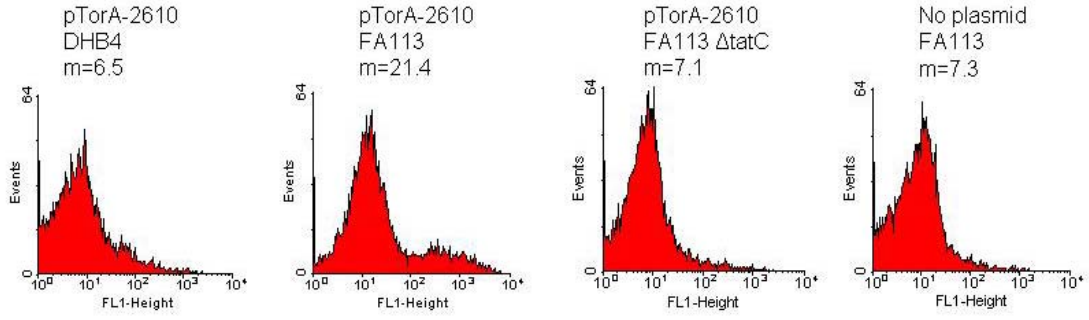

Figure 2.2. Flow cytometric analysis of ssTorA-26-10. *E. coli* strains DHB4, FA113 and FA113 ΔtatC are shown along with strain FA113 containing empty pBAD18 plasmid as a control. Cell populations with indicated fluorescence intensities are represented as histograms. m, geometric mean cell fluorescence intensity.

The amount of 26-10 scFv protein in the cytoplasmic and periplasmic fractions was evaluated by Western blotting following cell fractionation using the osmotic shock procedure [11, 32]. The 26-10 scFv protein was detected using an anti-hexahistidine monoclonal antibody that recognizes the C-terminal His6 tag. Consistent with the FACS analysis, the scFv protein could be detected in the periplasmic fraction from *E. coli* FA113 cells but not in *E. coli* DHB4 or FA113 Δ*tatC* (see Figure 2.3). In *E. coli* FA113 cells, the intensity of the antibody fragment band in the periplasmic fraction was only about 20-25% of the amount found in the cytoplasm indicating that export via Tat was inefficient. Accumulation of precursor protein in the cytoplasm is frequently observed when authentic Tat substrate proteins or fusions to Tat signal peptides are expressed from multicopy plasmids [37-40].

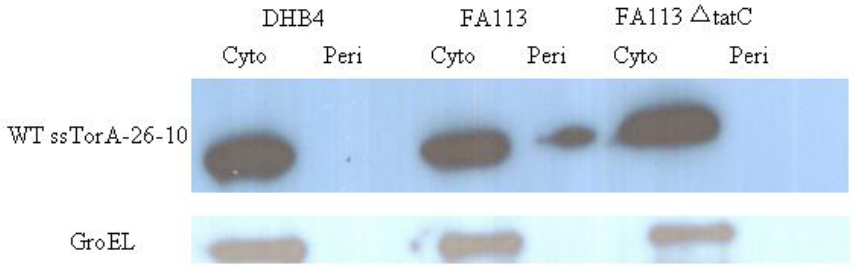

Figure 2.3. Western blot of cell fractions. Cell fractions from *E. coli* DHB4, FA113 and FA113 ΔtatC expressing ssTorA-26-10 scFv were performed by the cold osmotic shock procedure. Western blots were also probed with anti-GroEL antibody to show that cytoplasmic leakage did not occur. Amounts corresponding to equal number of cells were loaded in each lane, and results were reproducible when repeated with duplicate cultures.

2.3.2 Library screening

The fact that the amount of the 26-10 scFv in the periplasm can be monitored quantitatively by flow cytometry was exploited to screen for mutations that enhance export via the Tat pathway. By detecting the digoxigenin-BODIPY fluorescence, the amount of periplasmic scFv within each cell is determined directly in the absence of

reporter proteins that can otherwise dominate the solubility and folding kinetics of protein fusions [13, 41, 42].

The 26-10 scFv gene was subjected to random mutagenesis by error prone PCR [31] to give a library of $6x10^6$ independent transformants. Sequencing of six clones picked at random revealed an average of 13 nucleotide substitutions per gene which corresponds to a 1.3% error rate. The library cell population was labeled by incubation with 200 nM digoxigenin-BODIPY, and fluorescence was analyzed by flow cytometry. As expected, the library displayed a lower mean fluorescence than cells expressing 26-10 scFv since the majority of random mutations are deleterious for function [43] (see Figure 2.4). Clones exhibiting higher cell fluorescence were isolated by FACS. To facilitate the isolation of antibody variants that accumulate at a higher level in the periplasm, rather than mutants exhibiting higher hapten affinity, cell labeling was carried out using 200 nM digoxigenin-BODIPY, a concentration significantly higher than the 2.3 nM K_D of the 26-10 scFv antibody for digoxigenin [33]. 10^7 cells were sorted and the 0.3% most fluorescent events were collected. Because incubation in 5xPBS, which is required for outer membrane permeabilization, reduces the cell viability and thus increases the possibility that rare clones might be lost, the collected cells were lysed, and the ssTorA-26-10 scFv genes in that population were recovered by PCR. The PCR product was cloned back to the original vector and electroporated into *E. coli* XL1-Blue for plasmid propagation. Purified plasmid was obtained from this strain and used to transform *E. coli* FA113 to give rise to $6x10^4$ transformants. After growth of the pooled transformants in liquid TB media supplemented with antibiotic, cells were labeled and subjected to a second round of sorting as above. The cell fluorescence after each round of sorting increased, suggesting successful enrichment of more fluorescent cells (see Figure 2.4).

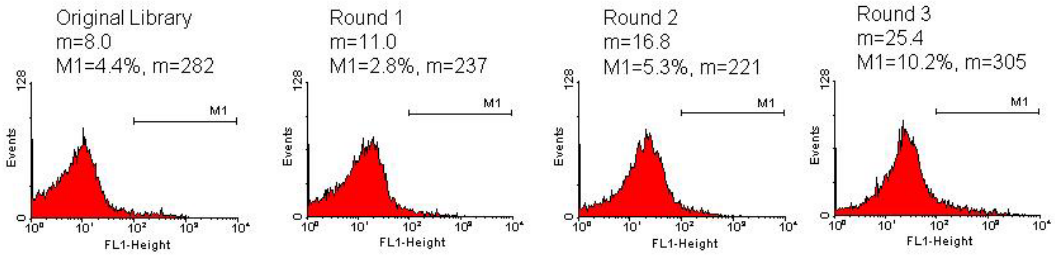

Figure 2.4. Enrichment of fluorescent clones by FACS. Histograms showing the mean fluorescent intensities of the original scFv library and the population of collected cells after each round of sorting. M1, marker indicating the percent of the total population in that region with the corresponding geometric mean fluorescence intensity, m.

Clones isolated after transformation from the third round were grown in 96-well plate format, and the fluorescence of each clonal population was determined. One clone C8, exhibiting the highest increase in fluorescence relative to the parental 26-10 scFv, was analyzed further. The fluorescence of cells expressing the C8 scFv was dependent on an oxidizing cytoplasm and was abolished in a Δ*tatC* strain background indicating that, as with the parental 26-10 scFv, both disulfide bond formation and a functional Tat pathway are required for export (see Figure 2.5). Consistent with the observed 2-fold increase in fluorescence observed with whole cells expressing the C8 scFv, about 40% of the C8 scFv protein was found in the periplasm compared to 20-25% for the parental antibody fragment (see Figure 2.6 and 2.3 for comparison).

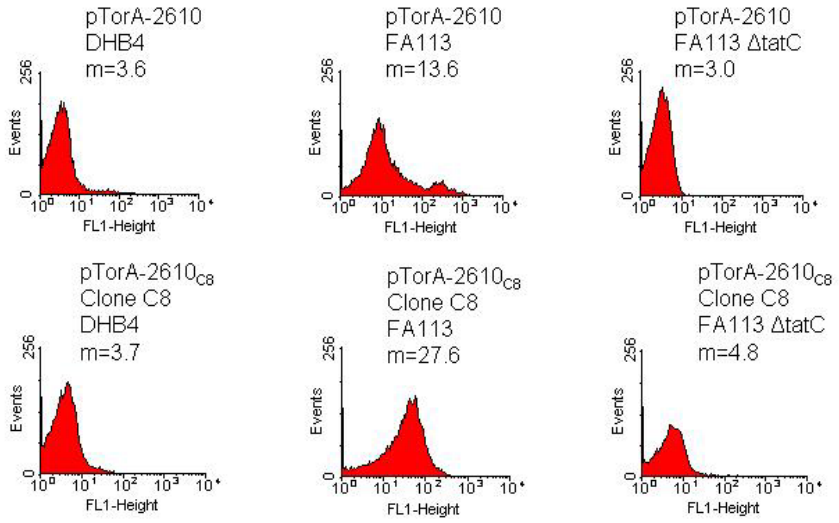

Figure 2.5. Fluorescence of C8 versus WT 26-10. Histograms of cells expressing ssTorA-2610$_{C8}$ isolated from the library or the WT ssTorA-26-10 scFv. m, geometric mean fluorescence intensity.

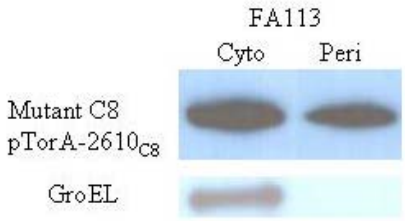

Figure 2.6. Subcellular localization of the C8 scFv. Western blot of cells expressing ssTorA-C8 26-10. Cells were fractionated using the osmotic shock procedure described above. Amounts corresponding to equal number of cells were loaded in each lane as in Figure 3.2 to allow direct comparison.

DNA sequencing revealed that the C8 clone contained 7 amino acid substitutions, 6 in the framework and 1 (L47 Leu→Pro, Kabbat numbering [44]) in CDR L2 (Figure 2.7). This residue does not make contact with the hapten. Also, none of the mutated positions corresponds to a second shell residue [45]. Second shell residues have side chains that are not directly involved in binding but whose orientation could affect the ligand from contacting the binding residues. In addition to C8, several lower

fluorescence clones (about 50% increase in whole cell fluorescence relative to the parental 26-10 scFv) were also isolated and were found to contain one or two amino acid substitutions. No consensus mutations were found among these clones.

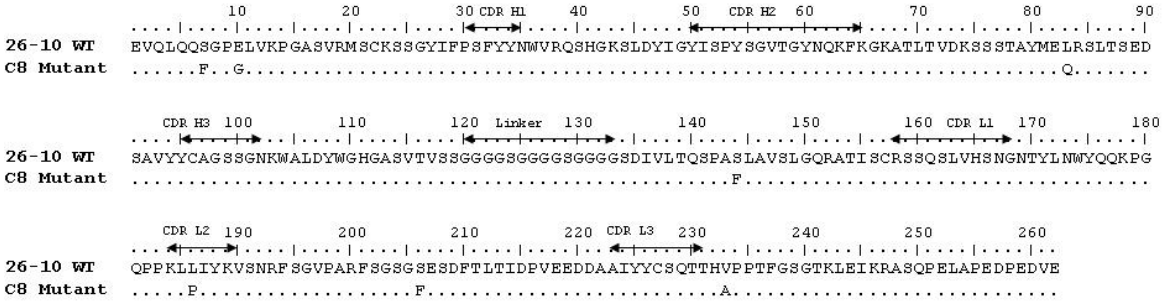

Figure 2.7. Alignment of C8 and WT scFvs. The positions of the amino acid substitutions in the C8 26-10 are shown. The residues that comprise the CDRs and linker are also indicated.

2.3.3 Expression and Characterization of the C8 scFv

To further confirm the increased export of the C8 scFv to the periplasm via the Tat pathway, purified protein was needed in order to characterize the mutant scFv. 1 L cultures were used to express enough C8 and WT scFvs for capture by immobilized metal chromatography and purification by gel filtration FPLC. The monomeric antibody fragments were purified to >95% homogeneity as determined by SDS-PAGE and quantified by bicinchoninic acid assay. The 26-10 and C8 scFvs were produced respectively at 0.13 mg/L/OD and 0.24 mg/L/OD. The increased yield of the purified C8 scFv is consistent with the approximately 2-fold higher level of periplasmic protein observed by Western blot analysis and with the increased cell fluorescence of C8 scFv-expressing cells (see Figures 2.3, 2.5 and 2.6).

Using the purified scFv protein, we sought to determine if the mutations in the C8 scFv affected the hapten binding kinetics or equilibrium dissociation constant. ELISA

assays indicate that the mutant C8 and WT 26-10 scFv have indistinguishable binding properties (see Figure 2.8). To further examine any subtle difference in hapten affinity, the hapten binding kinetics were determined quantitatively by surface plasmon resonance on a BIACore 3000 instrument. The association rate constant, k_a, dissociation rate constant k_d and the equilibrium dissociation rate constant K_D for C8 were 3.8 x 10^5 M^{-1} s^{-1}, 1.4 x 10^{-3} s^{-1} and 3.8 nM respectively and were statistically indistinguishable from the respective values for the parental 26-10 scFv (4.8 x 10^5 M^{-1} s^{-1}, 1.7 x 10^{-3} s^{-1} and 3.5 nM). The data above show that, consistent with the screening strategy that led to its isolation, the ability of C8 to confer increased cell fluorescence was not due to a higher hapten affinity but rather due to enhanced accumulation in the periplasm.

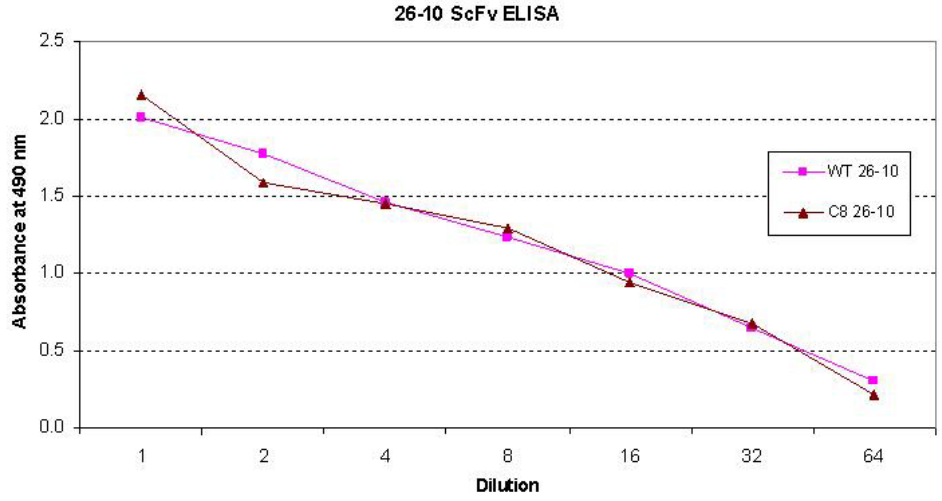

Figure 2.8. ELISA analysis of C8 and WT scFvs. Equal amounts of purified scFv were used to analyze binding to digoxigenin-BSA by ELISA.

Finally, a higher periplasmic yield of C8 was observed only when the protein was exported via the Tat pathway. Replacement of the ssTorA signal peptide with the *pelB* signal peptide directs export through the Sec pathway [25]. Precursor proteins competent for Sec export are maintained in a largely unfolded conformation within the cytoplasm in

contrast to proteins that are exported through Tat [46]. Since the formation of disulfide bonds inhibits Sec export, the accumulation of the scFv in the periplasm was examined in *E. coli* DHB4 which has a reducing cytoplasm and is isogenic to the oxidizing strain FA113 that had been used for Tat export. The fluorescence of cells expressing the ssPelB-26-10 C8 scFv fusion was indistinguishable from that of cells expressing the parental ssPelB-26-10 scFv (see Figure 2.9). Western blot analysis further revealed identical amounts of C8 and WT 26-10 scFvs in periplasmic fractions when exported by the Sec pathway (see Figure 2.10). Thus, the more efficient periplasmic localization of the C8 scFv occurs only when export is mediated by the Tat pathway.

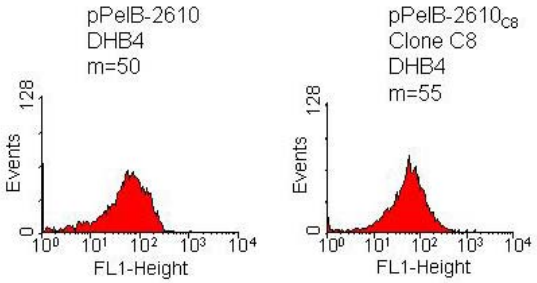

Figure 2.9. Sec export of C8 and WT scFvs**.** Fluorescence histograms of *E. coli* DHB4 expressing ssPelB-26-10 C8 or ssPelB-26-10 scFv which are exported via Sec.

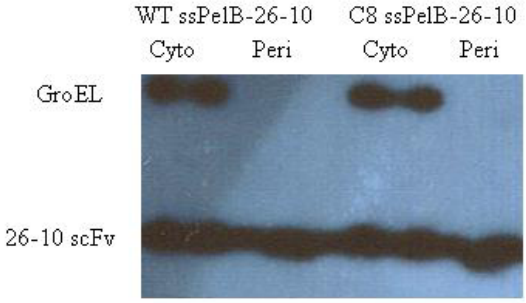

Figure 2.10. Western blot of cell fractions. Cell fractions obtained by cold osmotic shock of *E. coli* DHB4 showing Sec export of the mutant C8 and WT 26-10. The blot was probed with anti-GroEL to show that cytoplasmic protein leakage into the periplasmic fraction did not occur.

As shown in Figures 2.3 and 2.6, the higher amount of C8 scFv in the periplasm was accompanied by a reduction in the amount of protein that remained in the cytoplasm. The observed increase in the yield of periplasmic scFv can be due to one or more of the following factors: (i) increased stability of the C8 protein to degradation in the periplasm; (ii) more efficient targeting of the precursor polypeptide to the TatABC translocon; (iii) greater export competence. The finding that export of C8 and the 26-10 scFvs via Sec resulted in equal periplasmic yields suggests that the two proteins have similar *in vivo* stability. Targeting of the scFvs to the Tat translocon is mediated by the signal peptide. Since the mutations in the C8 scFv are all located in the mature protein they are less likely to influence the efficiency of translocon targeting [47]. The higher ratio of periplasmic to cytoplasmic C8 compared to the parental 26-10 scFv is consistent with the third possibility discussed above, namely more efficient export. Unfortunately, as is the case with many Tat precursor polypeptides, processing by signal peptidase in the periplasm is inefficient, and therefore the kinetics of export cannot be measured using pulse-chase techniques.

2.3.4 Folding Characteristics of the C8 scFv

We reasoned that the higher yield of periplasmic C8 may be related to the export competence of the protein in the cytoplasm which in turn is influenced by its folding kinetics. Therefore, in an effort to further elucidate the molecular basis for the improved export of C8 relative to the 26-10 scFv, we examined their guanidine HCl (GdnHCl)-induced equilibrium denaturation, folding kinetics and aggregation propensities. GdnHCl causes proteins to denature because it interferes with the hydrogen bonding network in water. By incubating a protein in different concentrations of GdnHCl, an equilibrium amount of folded versus unfolded protein is established. Fluorescence spectroscopy measures the amount of fluorescence from exposed tryptophan residues. As a higher

fraction of the protein molecules are unfolded, more tryptophan residues are available to fluoresce, therefore making the fluorescence proportional to the amount of unfolded protein.

The GdnHCl-induced denaturation transition was determined both for oxidized and for completely reduced scFv proteins (see Figure 2.11). Fluorescence spectroscopy revealed that increasing concentrations of GdnHCl caused a shift in the fluorescence emission maxima from 325 nm to 360 nm for the fully denatured proteins. Normalized transition curves were prepared based on the change in emission intensity at 360 nm. For the reduced C8 and 26-10 scFvs, the transition curves were essentially superimposable with midpoints at 2.75 M and 2.65 M, respectively. The transition midpoints for the oxidized scFvs exhibited a similar difference (3.60 M for C8 compared to 3.45 M for 26-10).

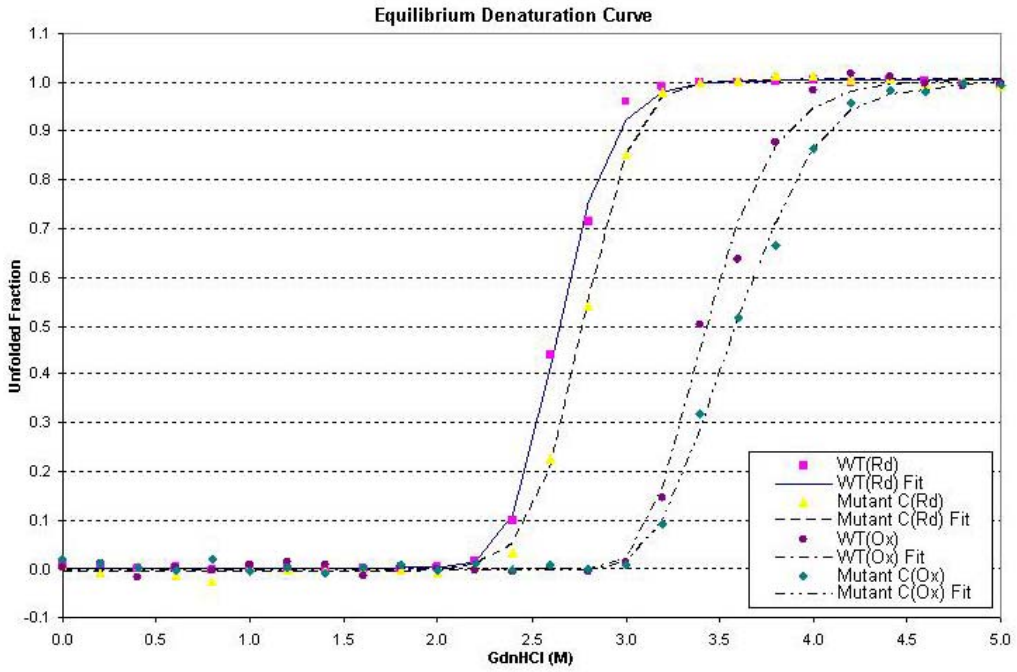

Figure 2.11. GdnHCl-induced denaturation equilibrium curves. Data is shown for both the reduced (Rd) and oxidized (Ox) mutant C8 and WT scFvs. The fitted lines are provided as visual aids.

Since the stability of the two scFvs was essentially identical, we decided to determine if their rates of folding were different. The folding kinetics were determined by diluting completely reduced and denatured proteins from 5 M GdnHCl to a final concentration of 0.5 M GdnHCl while following the change in fluorescence at 325 nm. Consistent with previous studies [36], the folding reaction exhibited three distinct phases. The first phase is complete during manual mixing (~10 sec) and accounts for 15-30% of the total amplitude of the fluorescence increase. The second phase follows first order kinetics and is completed within 180 seconds. Finally, a third slow phase was also observed. The third, slow phase has been shown to arise from peptidyl proline isomerization and exhibits a macroscopic rate of $<10^{-3}$ s^{-1} [24, 48]. Differences in the folding kinetics of C8 and 26-10 were observed during the second phase (see Figure 2.12). The increase in fluorescence data could be fit to a single exponential equation resulting in a rate constant of 0.063 s^{-1} for C8 and 0.025 s^{-1} for 26-10. In other words, the second critical phase in the folding of C8 occurs at a rate that is approximately 2.5 times higher than that of the parental 26-10 antibody. It should be noted that cysteine oxidation in *E. coli trxB gor* strains is relatively slow [35], and therefore the formation of the disulfide bonds in scFvs occurs at a late step, most likely prior to or concomitant with *cis-trans* proline isomerization.

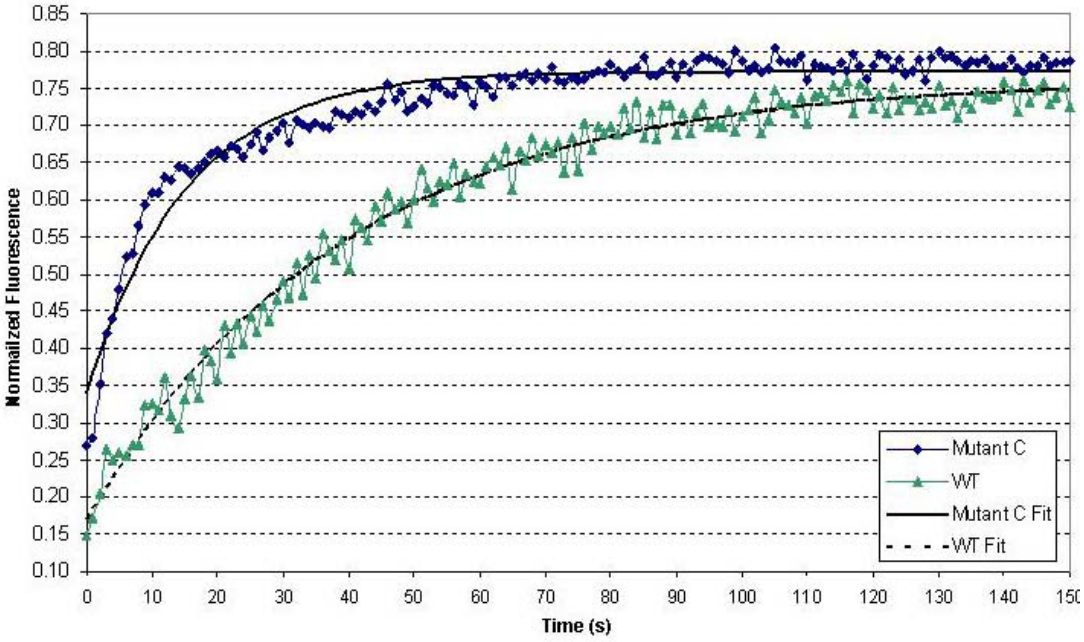

Figure 2.12. Renaturation of scFvs**.** By following the renaturation of the C8 and WT 26-10 scFvs, their kinetic rates of refolding were determined from the fitted curves.

Finally, we determined the aggregation propensity of the scFv fragments by monitoring the increase in light scattering following prolonged incubation in different concentrations of GdnHCl. Aggregation of protein while it is folding would lower the amount of competent protein available for export via the Tat pathway. As protein aggregates form, they scatter more light than soluble protein. Thus an increase in light scattering reflects an increase in aggregation. An increase in light scattering was observed following prolonged incubation in GdnHCl concentrations near the midpoint of the denaturation transition of both the reduced and the oxidized proteins (see Figure 2.13). In 2.5-3.0 M GdnHCl, the reduced C8 scFv exhibited approximately 30% lower light scattering than 26-10; a smaller difference was observed for the oxidized proteins, which showed a lower degree of aggregation.

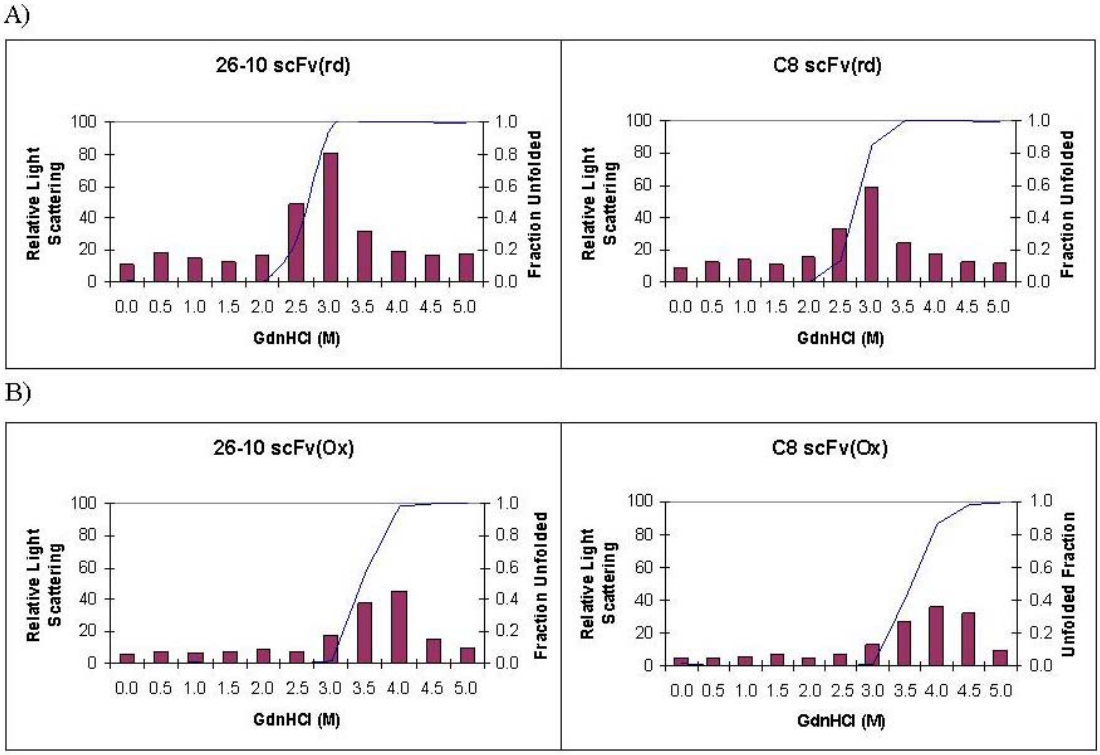

Figure 2.13. Aggregation of scFvs. Light scattering was used to determine the aggregation of both the A) reduced and B) oxidized C8 and WT scFvs upon prolonged incubation in GdnHCl. Bars show the light-scattering at the indicated GdnHCl concentrations. Lines represent the equilibrium denaturation curves as determined in Figure 2.11.

2.4 DISCUSSION

Export of the anti-digoxin 26-10 scFv antibody fragments via the Tat pathway by means of the ssTorA signal peptide can only occur in *E. coli* strains having an oxidizing cytoplasm that allow the formation of its two disulfide bonds. However, we found that the yield of Tat-exported 26-10 scFv in the periplasm was only 0.13 mg/L/OD which is considerably less than the yield obtained when export is mediated by a Sec signal peptide (approximately 1 mg/L/OD, unpublished laboratory result). A large fraction of the ssTorA-26-10 scFv is localized in the cytoplasm and is incompetent for export (see Figure 2.3). Accumulation of Tat precursor proteins in the cytoplasm is routinely observed not only with heterologous proteins but also with many native Tat substrates expressed at non-physiological levels [37-40]. We took advantage of a quantitative assay for monitoring the amount of periplasmic 26-10 scFv at the single cell level [28] to search for mutations in the mature protein that increase the export yield. A single clone, C8, which exhibited a 2-fold higher yield of periplasmic scFv and a proportional reduction in the amount of protein that remained in the cytoplasm, was isolated from a library of $6x10^6$ random mutants. Increased periplasmic accumulation was only observed when the protein was exported via the Tat pathway indicating that the mutations in the C8 scFv specifically affected a step important in protein translocation.

Comparison of the folding properties of the mutant and the parental scFvs revealed that the former was only marginally more stable to GdnHCl denaturation under either reducing or oxidizing conditions. However, kinetic experiments showed that for the reduced C8 scFv, the critical second phase in refolding is accelerated substantially. In addition, the C8 scFv exhibited a reduction in the aggregation of a folding intermediate populated at denaturant concentrations near the transition midpoint. The folding properties of the C8 scFv bear many similarities to those of the anti-β-galactosidase 13R4

scFv that was isolated earlier based on its ability to fold into an active conformation within the reducing environment of the *E. coli* cytoplasm [36, 49]. Similar to what we found for the anti-digoxigenin scFv mutant examined in the present study, the reduced 13R4 scFv exhibited a 3-fold acceleration in the intermediate refolding phase and a somewhat lower propensity to aggregate compared to its parental protein. Martineau and Betton suggested that the increase in the folding rate of the 13R4 scFv helps prevent off-pathway interactions that lead to aggregation [36]. A reduction in protein aggregation can also play a role in Tat export competence [13]. However, the export of scFvs is a complex process that also depends on disulfide bond formation, since only the oxidized protein can be translocated and ultimately, on the features that are recognized by the folding quality control mechanism of the Tat pathway.

Regardless of the precise mechanism, our data indicate that the rate of folding plays an important role in the export competence of scFv antibodies and possibly of other proteins. Along these lines we note that thioredoxin-1, a protein that exhibits very fast folding kinetics [9] when fused to ssTorA is exported with very high efficiency even when its expression is driven by very strong promoters (Masip and Georgiou, in preparation). In contrast, when thioredoxin is fused to a signal peptide that directs post-translational export via Sec, it is localized in the periplasm very inefficiently [9, 50]. Thus, it is tempting to speculate that Tat export hinges on folding properties that are diametrically opposite to those required for post-translational export through the Sec translocon. If that is the case, genetic screens for improved Tat export could be useful for selecting mutants of soluble proteins exhibiting faster folding. Apart from its practical ramifications for protein expression, the isolation of faster folding proteins can provide insights into the nature of the rate limiting steps in folding.

Finally, the C8 scFv mutant has an increased yield of approximately two-fold over the parental 26-10 scFv. Therefore, the genetic selection for a mutant scFv with improved periplasmic yield through the Tat pathway was successful. However, since this improved yield is still lower than that achieved when export is targeted to the Sec pathway, additional studies must be performed to further increase the yield and allow the Tat pathway to become a viable mechanism for protein production in biotechnology. Such experiments are detailed in the next chapter and use the mutant C8 mutant scFv as a protein reporter.

2.5 REFERENCES

1. Schatz, G. and B. Dobberstein, *Common principles of protein translocation across membranes.* Science, 1996. **271**(5255): p. 1519-26.
2. Kim, J. and D.A. Kendall, *Sec-dependent protein export and the involvement of the molecular chaperone SecB.* Cell Stress Chaperones, 2000. **5**(4): p. 267-75.
3. Breyton, C., et al., *Three-dimensional structure of the bacterial protein-translocation complex SecYEG.* Nature, 2002. **418**(6898): p. 662-5.
4. Connolly, T. and R. Gilmore, *The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide.* Cell, 1989. **57**(4): p. 599-610.
5. Nakonechny, W.S. and C.M. Teschke, *GroEL and GroES control of substrate flux in the in vivo folding pathway of phage P22 coat protein.* J Biol Chem, 1998. **273**(42): p. 27236-44.
6. Doyle, S.M., et al., *Rapid unfolding of a domain populates an aggregation-prone intermediate that can be recognized by GroEL.* J Mol Biol, 2003. **332**(4): p. 937-51.
7. Randall, L.L., et al., *Sites of interaction between SecA and the chaperone SecB, two proteins involved in export.* Protein Sci, 2004. **13**(4): p. 1124-33.
8. Randall, L.L. and S.J. Hardy, *SecB, one small chaperone in the complex milieu of the cell.* Cell Mol Life Sci, 2002. **59**(10): p. 1617-23.
9. Huber, D., et al., *A selection for mutants that interfere with folding of Escherichia coli thioredoxin-1 in vivo.* Proc Natl Acad Sci U S A, 2005. **102**(52): p. 18872-7.
10. Lee, P.A., D. Tullman-Ercek, and G. Georgiou, *The Bacterial Twin-Arginine Translocation Pathway.* Annu Rev Microbiol, 2006.
11. DeLisa, M.P., D. Tullman, and G. Georgiou, *Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6115-20.
12. Richter, S. and T. Bruser, *Targeting of unfolded PhoA to the TAT translocon of Escherichia coli.* J Biol Chem, 2005. **280**(52): p. 42723-30.
13. Fisher, A.C., W. Kim, and M.P. DeLisa, *Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway.* Protein Sci, 2006. **15**(3): p. 449-58.
14. Halbig, D., et al., *The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding.* Eur J Biochem, 1999. **263**(2): p. 543-51.
15. Rodrigue, A., et al., *Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway.* J Biol Chem, 1999. **274**(19): p. 13223-8.

16. Thomas, J.D., et al., *Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli.* Mol Microbiol, 2001. **39**(1): p. 47-53.
17. Clark, S.A. and S.M. Theg, *A folded protein can be transported across the chloroplast envelope and thylakoid membranes.* Mol Biol Cell, 1997. **8**(5): p. 923-34.
18. Hynds, P.J., D. Robinson, and C. Robinson, *The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane.* J Biol Chem, 1998. **273**(52): p. 34868-74.
19. Robinson, C. and A. Bolhuis, *Protein targeting by the twin-arginine translocation pathway.* Nat Rev Mol Cell Biol, 2001. **2**(5): p. 350-6.
20. Li, S.Y., B.Y. Chang, and S.C. Lin, *Coexpression of TorD enhances the transport of GFP via the TAT pathway.* J Biotechnol, 2006. **122**(4): p. 412-21.
21. Hatzixanthis, K., et al., *Signal peptide-chaperone interactions on the twin-arginine protein transport pathway.* Proc Natl Acad Sci U S A, 2005. **102**(24): p. 8460-5.
22. Holt, L.J., et al., *Domain antibodies: proteins for therapy.* Trends Biotechnol, 2003. **21**(11): p. 484-90.
23. Hoyer, W., K. Ramm, and A. Pluckthun, *A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments.* Biophys Chem, 2002. **96**(2-3): p. 273-84.
24. Jager, M. and A. Pluckthun, *Folding and assembly of an antibody Fv fragment, a heterodimer stabilized by antigen.* J Mol Biol, 1999. **285**(5): p. 2005-19.
25. Daugherty, P.S., et al., *Antibody affinity maturation using bacterial surface display.* Protein Eng, 1998. **11**(9): p. 825-32.
26. Huston, J.S., et al., *Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli.* Proc Natl Acad Sci U S A, 1988. **85**(16): p. 5879-83.
27. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.
28. Chen, G., et al., *Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS).* Nat Biotechnol, 2001. **19**(6): p. 537-42.
29. Emmerson, P.J., et al., *Synthesis and characterization of 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (BODIPY)-labeled fluorescent ligands for the mu opioid receptor.* Biochem Pharmacol, 1997. **54**(12): p. 1315-22.
30. Hayhurst, A., et al., *Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis.* J Immunol Methods, 2003. **276**(1-2): p. 185-96.
31. Fromant, M., S. Blanquet, and P. Plateau, *Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction.* Anal Biochem, 1995. **224**(1): p. 347-53.
32. Levy, R., et al., *Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones.* Protein Expr Purif, 2001. **23**(2): p. 338-47.

33. Chen, G., et al., *In vitro scanning saturation mutagenesis of all the specificity determining residues in an antibody binding site.* Protein Eng, 1999. **12**(4): p. 349-56.
34. Pace, C.N., *Measuring and increasing protein stability.* Trends Biotechnol, 1990. **8**(4): p. 93-8.
35. Bessette, P.H., et al., *Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13703-8.
36. Martineau, P. and J.M. Betton, *In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm.* J Mol Biol, 1999. **292**(4): p. 921-9.
37. Chanal, A., C.L. Santini, and L.F. Wu, *Specific inhibition of the translocation of a subset of Escherichia coli TAT substrates by the TorA signal peptide.* J Mol Biol, 2003. **327**(3): p. 563-70.
38. DeLisa, M.P., et al., *Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway.* J Bacteriol, 2004. **186**(2): p. 366-73.
39. Mikhaleva, N.I., et al., *Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in Escherichia coli.* FEBS Lett, 1999. **463**(3): p. 331-5.
40. Stanley, N.R., et al., *Behaviour of topological marker proteins targeted to the Tat protein transport pathway.* Mol Microbiol, 2002. **43**(4): p. 1005-21.
41. Harvey, B.R., et al., *Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries.* Proc Natl Acad Sci U S A, 2004. **101**(25): p. 9193-8.
42. Cabantous, S., et al., *Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from Mycobacterium tuberculosis.* J Struct Funct Genomics, 2005. **6**(2-3): p. 113-9.
43. Daugherty, P.S., et al., *Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2029-34.
44. Martin, A., *Accessing the Kabat Antibody Sequence Database by Computer.* PROTEINS: Structure, Function and Genetics, 1996. **25**: p. 130-133.
45. Jeffrey, P.D., et al., *26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity.* Proc Natl Acad Sci U S A, 1993. **90**(21): p. 10310-4.
46. Knoblauch, N.T., et al., *Substrate specificity of the SecB chaperone.* J Biol Chem, 1999. **274**(48): p. 34219-25.
47. Alami, M., et al., *Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli.* Mol Cell, 2003. **12**(4): p. 937-46.
48. Jager, M., P. Gehrig, and A. Pluckthun, *The scFv fragment of the antibody hu4D5-8: evidence for early premature domain interaction in refolding.* J Mol Biol, 2001. **305**(5): p. 1111-29.
49. Martineau, P., P. Jones, and G. Winter, *Expression of an antibody fragment at high levels in the bacterial cytoplasm.* J Mol Biol, 1998. **280**(1): p. 117-27.

50. Schierle, C.F., et al., *The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway.* J Bacteriol, 2003. **185**(19): p. 5706-13.

Chapter 3

Increasing Tat Export Efficiency

3.1 INTRODUCTION

As discussed in Chapter 1, the Tat pathway exhibits a quality control feature that results in the export of only fully and correctly folded proteins to the periplasm [1]. However, secretion via Tat is not widely utilized since its efficiency is much lower than that of the Sec export pathway [2]. In addition, protein yields obtained through the Tat pathway do not approach those obtained from Sec export which, for antibody fragments, typically exceed 1 g/L [3, 4] although this is from fermentation studies and may not be typical. As indicated in the last chapter, the yield of the C8 scFv was only 0.24 mg/L/OD. This lower yield may be due to the fact that the export machinery can be saturated by overexpressed target proteins. Saturation of the Tat machinery has been reported even for native Tat export proteins [5-8]. In some instances, the saturation of export can be partially relieved by co-expression of other proteins [6, 9]. However, it may also be necessary to optimize growth conditions to increase expression and/or engineer a better performing strain.

Recent studies have indicated that overexpression of TatABC results in improved secretion [5, 10]. However, since these are all membrane proteins, expression beyond a certain level may adversely affect cell growth and will require determining the optimal levels experimentally [11]. TatA is believed to form the protein conducting pore [12, 13], and therefore overexpression of TatA by itself may allow more pores to form by increasing the availability of TatA subunits for oligomerization. Since the Tat pathway

uses the proton motive force (PMF) to export proteins, recent work from our laboratory has shown that overexpression of PspA, a phage shock protein associated with the maintenance of the PMF, partially relieves the saturation of the Tat pathway [6]. DsbC is a disulfide isomerase which functions to help periplasmic proteins form correct disulfide bonds. Thus, for disulfided proteins, overexpression of a cytoplasmic version of DsbC might help in correct disulfide bond formation [14] allowing more protein to be competent for export. Finally, TorD has been shown to be a chaperone which protects the Tat signal peptide from degradation in the cytoplasm before export [15]. This protection may allow more precursor proteins to be targeted to the Tat pathway, and overexpression of TorD has been shown to increase Tat export [9].

Environmental factors such as temperature, pH, and the availability of nutrients play an important role in gene expression of both native and recombinant proteins in *E. coli* [16-20]. In addition, proteins tend to fold incorrectly or become insoluble at physiological temperatures and therefore, expression typically is carried out at lower temperatures [21, 22]. Since the Tat pathway requires proteins to be soluble and correctly folded, the growth temperature at induction could play a role in increasing the amount of export through the pathway. Media and pH must also be optimized to not only allow for maximum cell growth but also to ensure expression is maximized, i.e., translation is not limited due to a lack of availability of amino acids. pH also can affect the PMF by influencing the potential difference across the bacterial membrane [23].

To investigate the effect of growth conditions on Tat export, the C8 ssTorA-26-10 scFv was used as a reporter. The C8 scFv can only be translocated into the periplasm when the gene fusion is expressed in *E. coli* strains that have an oxidizing cytoplasm, such a FA113. This strain contains mutations in the *trxB* and *gor* genes which allow the formation of the two disulfide bonds in scFv that are necessary to stabilize the protein [1,

24, 25]. As described in the previous chapter, the *trxB* and *gor* genes of *E. coli* maintain the reducing nature of the cytoplasm. Thioredoxins and glutaredoxins can catalyze the formation of disulfide bonds in their oxidized state but are kept in a reduced state by thioredoxin reductase, encoded by *trxB*, and glutathione oxidoreductase, encoded by the *gor* gene. However, due to the accumulation of these oxidized proteins in the cytoplasm, the growth rate of these mutant oxidizing strains is slower compared to the WT parental strains [24]. Different strains have been created with additional mutations in the *gor* and *trx* genes to create an oxidizing cytoplasm while attempting to reduce the amount of oxidized protein accumulation. For example, strains DR602 and DR611 are also lacking the *trxA* and *trxC* genes encoding for thioredoxin and thioredoxin2, respectively. Expressing the C8 scFv in these strains may allow the formation of the disulfide bonds while providing better growth characteristics than the FA113 strain used previously.

In this study, we sought to analyze different methods to increase the efficiency of Tat export using the 26-10 and the C8 scFv mutant as protein reporters. The export yield of both proteins can be conveniently measured by FACS by monitoring the cell fluorescence following permeabilization of the outer membrane and labeling with digoxigenin-BODIPY [26]. An increase in Tat export was observed under certain conditions. Combined with the two-fold higher yield of the C8 scFv, the overall yield of the anti-digoxin antibody fragment was increased by about 5-fold relative to the 26-10 scFv. The effect of growth media, temperature and pH changes were also investigated for their effect on Tat export. However, no changes in these variables had a positive impact. Finally, five different oxidizing cytoplasm strains were also used to express C8 26-10. Although some of these produced fluorescence approaching the FA113 strain used previously, the cell densities achieved with the other strains were less than that of FA113, translating into an overall lower yield of 26-10.

3.2 MATERIALS AND METHODS

3.2.1 Strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table 3.1. *E. coli* XL1-Blue cells were used for plasmid propagation.

To construct the pPspA plasmid, primers PspA.f and PspA.r were used to amplify the *pspA* gene from p99PspA by PCR. The PCR fragment and empty pBAD18 plasmid were digested by *Xba*I and *Sac*I and ligated forming the pPspA plasmid.

Using primers TatA.f and TatA.r and genomic DNA as template, PCR was carried out to clone the *tatA* gene. The DNA fragment containing the *tatA* gene was digested by *Xba*I and *Sac*I and ligated into similarly digested empty pBAD18 to yield pTatA.

pTorD pBAD33 was constructed by PCR with primers TorD.f and TorD.r and p99TorD pTrc99A as template. This cloned fragment and empty pBAD33 were digested with *Xba*I and *Sac*I, and upon ligation, pTorD was created.

The C8 scFv gene in pTorA-2610$_{C8}$, was used as template for PCR amplification with primers C8trc.f and C8trc.r. The resulting PCR product was digested with *Nhe*I and *Xho*I and ligated at the respective sites of pTrc99A. Ligation resulted in the creation of plasmid p99TorA-2610$_{C8}$. The p99TorA-2610$_{C8}$ plasmid was created to ensure plasmid compatibility with the pBAD plasmids used in the co-expression studies.

Table 3.1. Strains, plasmids, and primers used in this study

E. coli Strain or Plasmid	Relevant Genotype or Features	Source
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q} *ZΔM15* Tn*10* (Tet^{r})]	Lab collection
DHB4	MC4100 *phoR* Δ(*phoA*) *Pvu*II Δ(*malF)3* F' [*lacI*^{q}*ZYA -pro*]	Lab collection
FA113	DHB4 *trxB gor552* Tn*10*tet^{r} *ahpC**	J. Beckwith
FA113 ΔtatC	DHB4 *trxB gor552* Tn*10*tet^{r} *ahpC** *tatC::spec*	Lab collection
DR600	DHB3 Δ*trxB (araC Para-trxB) gor::km katG::*Tn*10tet*^{r} *ahpC*(P166L)	J. Beckwith
DR602	DHB3 Δ*trxA* Δ*trxC trxB::km gor522* Tn*10*tet^{r} *ahpC*(T104P)	J. Beckwith
DR611	DHB3 Δ*trxA* Δ*trxC trxB::km gor522* Tn*10*tet^{r} *ahpC*(T104P/G141C)	J. Beckwith
SMG96	DHB4 Δ*trxB* Δ*gor ahpC**	J. Beckwith
MJF256.10	DHB4 Δ*gshA::km* Δ*trxB::cm ahpC*(V164G)	J. Beckwith
pBAD18	*araBAD* promoter, pBR322 *ori*, cm^{r}	Ref. [27]
pBAD33	*araBAD* promoter, pBR322 *ori*, cm^{r}	Ref. [27]
pTrc99A	*trc* promoter, ColE1 *ori*, cm^{r}	Amersham Biosciences
p99PspA	PspA in pTrc99A	Lab collection
pPspA	PspA in pBAD18	This study
pEXT-ABC	*tac* promoter, *lacIq*, *tatA*, *tatB*, *tatC*, Kan^{r}	Ref. [28]
pΔssDsbC	ΔssDsbC in pBAD33	Lab collection
pTatA	TatA in pBAD18	This study
p99TorD	TorD in pTrc99A	Lab collection
pTorD	TorD in pBAD33	This study
p99TorA-2610	ssTorA-26-10 in pTrc99A	Lab collection
p99TorA-2610$_{C8}$	ssTorA-26-10 mutant C8 in pTrc99A	This study
pTorA-2610$_{C8}$	ssTorA-26-10 mutant C8 in pBAD18	Chapter 3
Primers		
PspA.f	GCGATGGAGCTCTTAAAGAGGAGAAAAGGTCATGGGTATTTTTTCTCGCT	
PspA.r	GCGATGTCTAGAGCATTATTGATTGTCTTGCTTCAT	
TatA.f	GCGATGGAGCTCTTAAAGAGGAGAAAAGGTCATGGGTGGTATCAGTATTTG	
TatA.r	GCGATGTCTAGAGCATTACACCTGCTCTTTATCG	
TorD.f	GCATGAGCTCTTAAAGAGGAGAAAGGTCGGTACCATGACCACGCTGACAG	
TorD.r	GCATTCTAGATTATCTGTTTTGGTGGTCG	
C8trc.f	GCGGCGACTGACGCTGCTAGCGAAGTT	
C8trc.r	AGCCCGTTTGATCTCGAG	

3.2.3 Growth conditions, expression and flow cytometry

Unless otherwise noted, cells were grown at 37^{o}C on LB media with 25 μg/ml chloramphenicol and/or 100 μg/ml ampicillian and/or 25 μg/ml kanamycin, as appropriate.

For flow cytometric analysis, cells containing the appropriate plasmids were grown overnight in TB media, unless noted otherwise, with the appropriate antibiotics. 250 μl of overnight culture was used to inoculate 10 ml of fresh media, and after incubation with shaking for 1 hour at 37^{o}C, protein expression was induced by the addition of arabinose and/or IPTG to a final concentration of 0.2% and 0.1 mM, respectively. The culture was then grown at 30^{o}C, unless noted otherwise, for four hours. 100 μl of the culture was harvested and labeled as described in Chapter 2. Depending upon the cell density, 3 to 10 μl of each labeled sample was added to 1 ml of 5xPBS in a sieve tube before being analyzed on a Becton-Dickinson FACSort flow cytometer with the same settings as described in Chapter 2.

3.3 RESULTS

3.3.1 Effect of co-expression of various genes on the periplasmic yield of the 26-10 scFv

3.3.1.A. PspA

PspA is a protein that is equally distributed between the inner membrane fraction and cytoplasm [29]. Earlier studies have indicated that PspA helps the cell manage membrane-related stress, such as osmotic shock or the blockage of the Sec pore, by assisting in the maintenance of the proton motive force (PMF) [8, 30, 31]. Kleerebezem *et al*. have shown that transport of proteins through the Sec pathway is less efficient in the absence of PspA while its expression increases the efficiency of Sec export [32]. The Tat pathway uses the PMF to drive the translocation of proteins across the membrane and recent studies by the Georgiou laboratory have shown that overexpression of PspA partially relieves the saturation of the Tat pathway observed during the export of a ssTorA-GFP fusion [6]. Therefore, we sought to determine the effect of co-expression of PspA on the periplasmic yield of the 26-10 scFv exported via the ssTorA signal peptide.

The *pspA* gene was cloned into pBAD18 to allow arabinose inducible, high level expression and compatibility with the p99TorA-2610 and p99TorA-2610_{C8} plasmids. pTrc99A was used for the 26-10 and C8 scFvs since it allows a high level of inducible expression with IPTG. The pPspA plasmid was used to transform *E. coli* FA113 cells containing either of the p99TorA-2610 and p99TorA-2610_{C8} plasmids that encode, respectively, the 26-10 and the C8 scFvs. Following growth and induction of PspA and scFv protein synthesis with arabinose and IPTG, the periplasmic yield of scFv was analyzed by flow cytometry. As seen in Figures 3.1B and 3.2B, the fluorescence signals from the expression of the 26-10 and the C8 scFvs were increased 1.9-fold and 1.5-fold,

respectively. An equivalent fold increase was observed for each scFv when the experiment was duplicated.

3.3.1.B. ΔssDsbC

DsbC is an isomerase that catalyzes the rearrangement of incorrectly formed disulfide bonds and is present in the periplasm of bacteria [33]. DsbC has also been found to have chaperone activity as well in that it assists in the folding of proteins that do not contain disulfide bonds [34]. Overexpression of DsbC has been shown to increase the isomerization and yield of non-native disulfide-containing proteins such as human nerve growth factor [35], tissue plasminogen activator [36] and scFvs [37].

By removing the putative signal peptide from DsbC, the enzyme is confined to the cytoplasm where it can catalyze correct disulfide bond formation in *trxB gor* mutant strains having an oxidizing cytoplasm. By facilitating disulfide bond formation in the cytoplasm, DsbC can increase the amount of protein that is competent for export [14]. Since the cytoplasm is oxidizing in the FA113 strain, it would allow DsbC to function as if it were normally in the periplasm.

The pΔssDsbC plasmid was used to transform FA113 cells expressing the 26-10 and C8 scFv fusions as above. Cell growth and induction were carried out as for cells co-expressing PspA (see above). As seen in Figures 3.1C and 3.2C, the fluorescence signals of the 26-10 and C8 scFvs were increased 1.6-fold and 1.4-fold, respectively. The increase in fluorescence of the cells expressing the C8 scFv is slightly lower than that of the 26-10 scFv. DsbC is normally regenerated to its active state by transfer of electrons from the periplasmic side of DsbD, located in the inner membrane. Since this regeneration mechanism is not available to cytoplasmic DsbC, recycling of the enzyme to its reduced state must be mediated by other mechanisms and may be the rate limiting step in disulfide bond formation.

3.3.1.C. TorD

TorD is a chaperone which protects the ssTorA signal peptide in the native ssTorA-TorA Tat substrate from degradation in the cytoplasm before export [15]. In addition, TorD also interacts with the TorA protein to promote its stability at higher growth temperatures [38]. By binding to immature proteins, TorD has been shown to prevent the premature binding of precursor proteins to the Tat machinery before proper cofactor insertion [39]. Recently, Li *et al.* found that the expression of TorD increases the export of ssTorA-GFP three-fold [9].

We examined whether co-expression of TorD can also facilitate the export of a protein with a more complex folding pattern, such as a scFv antibody. The pTorD plasmid was constructed by amplifying *torD* by PCR from p99TorD and ligating the digested DNA fragment into the *Xba*I and *Sac*I sites of pBAD33. pTorD was transformed into FA113 cells containing the scFv fusions. Flow cytometry revealed that co-expression of TorD had a slight effect on cell fluorescence and resulted in a 40% and 30% increase for the 26-10 and C8 scFvs, respectively (see Figures 3.1D and 3.2D). This small increase for both scFvs may be due to the protection of the ssTorA peptide by TorD while awaiting export. However, the increase in export of the scFvs was lower than the three-fold increase reported earlier for the ssTorA-GFP protein reporter. This may be attributed to differences in folding between the scFv and GFP, in that the scFv folding or disulfide formation may sterically interfere with the binding of TorD. Perhaps, the lower amount of protection to the ssTorA signal peptide conferred by TorD results in proteolysis of the signal peptide at some stages of the folding pathway.

3.3.1.D. TatA

As discussed in Chapter 1, TatA is expressed by cells at approximately 20 times the level of the other two essential Tat pore component proteins, TatB and TatC [40], indicating that TatA most likely forms the majority of the Tat pore [13, 41, 42]. Electron microscopy of a Tat complex revealed the formation of pore-like structures comprised mainly of TatA. These pore-like structures appeared to be occluded on the cytoplasmic side of the inner membrane consistent with the proposed role of TatA as the main component of the gated Tat pore [12]. Overexpression of TatA may allow more pores to form by increasing the TatA subunits available for oligomerization and, in turn, increase the flux of protein exported across the membrane. The *tatA* gene was cloned into pBAD18 to allow a high level of expression. Following transformation of pTatA into the FA113 cells containing the two anti-digoxin scFvs, export was evaluated by flow cytometry as above. Once again, only a slight increase in cell fluorescence was observed (see Figures 3.1E and 3.2E). Flow cytometric analysis of duplicate cultures for the 26-10 and C8 scFvs produced statistically similar increases in fluorescence, respectively.

3.3.1.E. Co-expression of the TatABC operon

Thus, recent studies have indicated that overexpression of TatABC can improve secretion through the Tat pathway [5, 10, 43]. However, since these are all membrane proteins, overexpression may adversely affect cell growth [11], reducing the overall protein yield even though export is increased.

The TatABC proteins were expressed as an operon under the control of the *tac* promoter from the pEXT-ABC plasmid [28]. Upon induction of protein synthesis, the expression of TatA, TatB, and TatC increases 10-fold relative to the basal level observed in cells where TatABC synthesis is directed from its native promoter in the chromosome [28]. Cell growth and protein synthesis was carried out as above except that IPTG was

used to induce both TatABC and scFv synthesis. Figures 3.1F and 3.2F show the fluorescence histograms of cells co-expressing the Tat pore proteins. In this case, a significant increase of 2.1- and 2.4-fold, respectively, in the fluorescence of the cells expressing the 26-10 and C8 scFvs, was observed by flow cytometry. Overexpression of TatABC in both cases show that the number of Tat translocases limited the efficiency of export. Additional experiments support the reproducibility of the results for each scFv.

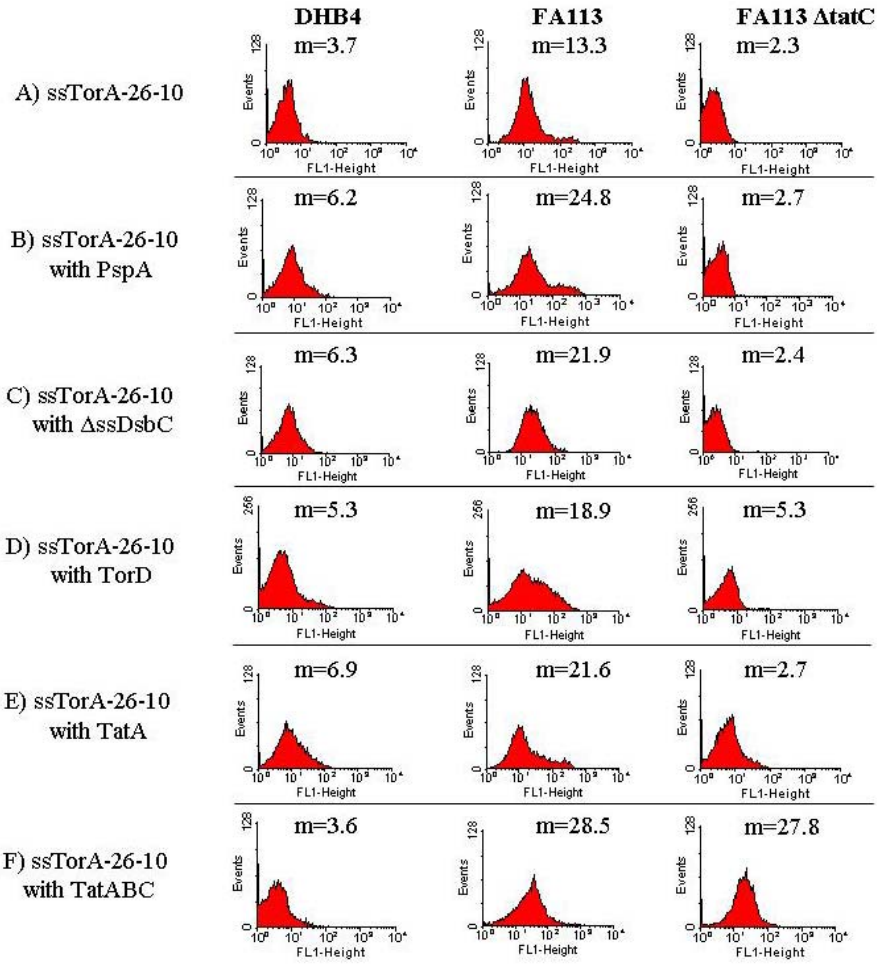

Figure 3.1. Effect of co-expression of various proteins on the export of 26-10 as monitored by flow cytometry. *E. coli* strains DHB4, FA113 and FA113 ΔtatC were transformed with p99TorA-2610 together with plasmids expressing the appropriate protein factor. A) Expression of the ssTorA-26-10 control. ssTorA-26-10 co-expressed with B) PspA, C) ΔssDsbC, D) TorD, E) TatA and F) TatABC. m, geometric mean fluorescence intensity of the histogram.

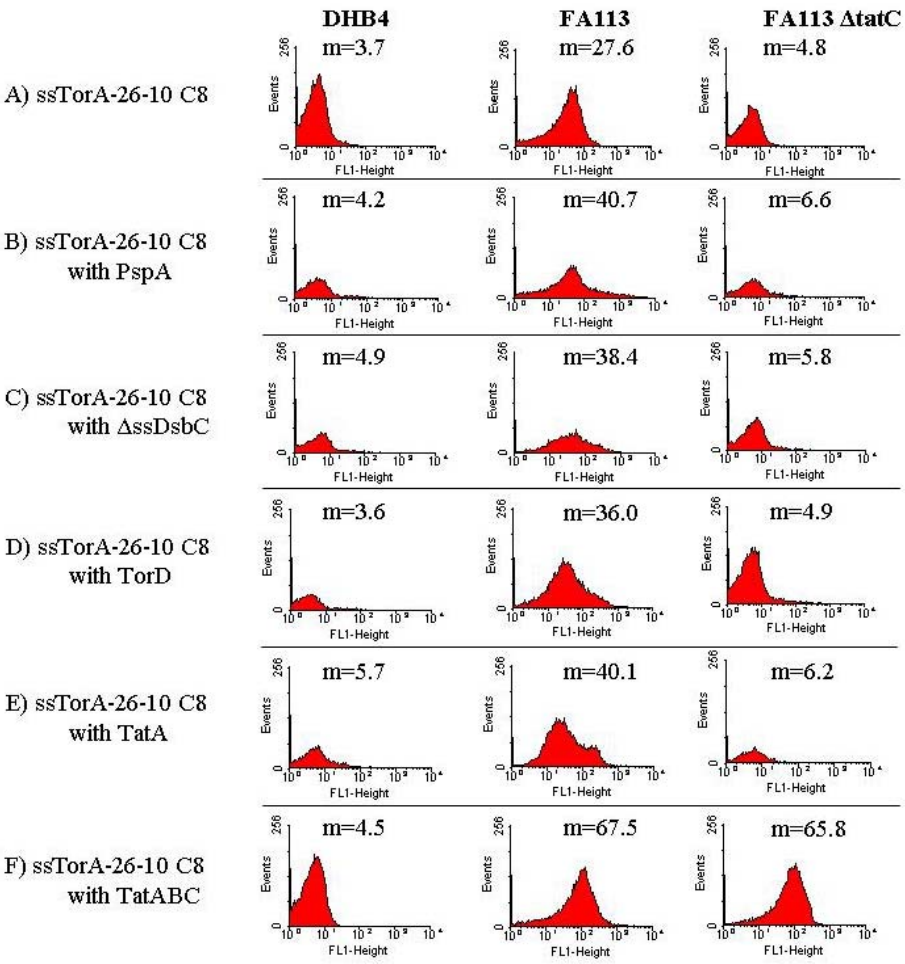

Figure 3.2. Effect of co-expression of various proteins on the export of C8 scFv as monitored by flow cytometry. *E. coli* strains DHB4, FA113 and FA113 ΔtatC were transformed with p99TorA-2610$_{C8}$ together with plasmids expressing the appropriate protein factor. A) Expression of the ssTorA-26-10 mutant C8 control. ssTorA-26-10 mutant C8 co-expressed with B) PspA, C) ΔssDsbC, D) TorD, E) TatA and F) TatABC. m, geometric mean fluorescence intensity of the histogram.

3.3.2 Effect of growth conditions on Tat export of scFv antibodies

Environmental factors such as temperature, pH and nutrient composition play an important role in determining the growth rate and in protein expression. These factors not only influence the maximum cell density a culture can obtain, but can also limit recombinant protein expression by preferentially increasing native gene and protein expression over recombinant proteins or not supplying enough amino acids in the media for high levels of translation [19, 20]. A higher growth temperature allows for a faster growth rate. On the other hand, proteins tend to fold incorrectly or become insoluble at supra-optimal or even physiological temperatures so induction of expression typically is carried out at lower temperatures [21, 22]. This is especially significant for Tat export since the Tat pathway requires proteins to be soluble and correctly folded in order to be competent for export. A high induction temperature could play a role in protein misfolding or aggregation, decreasing the amount of export through the pathway. In addition, growth at lower temperatures increases the permeability of the outer membrane [44] and is preferred for flow cytometric analysis because it allows better penetration of the fluorescent probe.

We sought to examine the effects of the growth and induction temperatures on the fluorescence of cells expressing the C8 scFv antibody exported via the ssTorA signal peptide. The C8 scFv was used as a reporter protein because it confers higher cell fluorescence and therefore the signal-to-noise ratio in these experiments is maximized relative to the export of the 26-10 scFv. Different growth media may allow achievement of higher cell densities, the cell to maintain the oxidizing cytoplasm more efficiently or increase the efficiency of transcription, translation or folding of the C8 scFv. The effect of temperature was also examined.

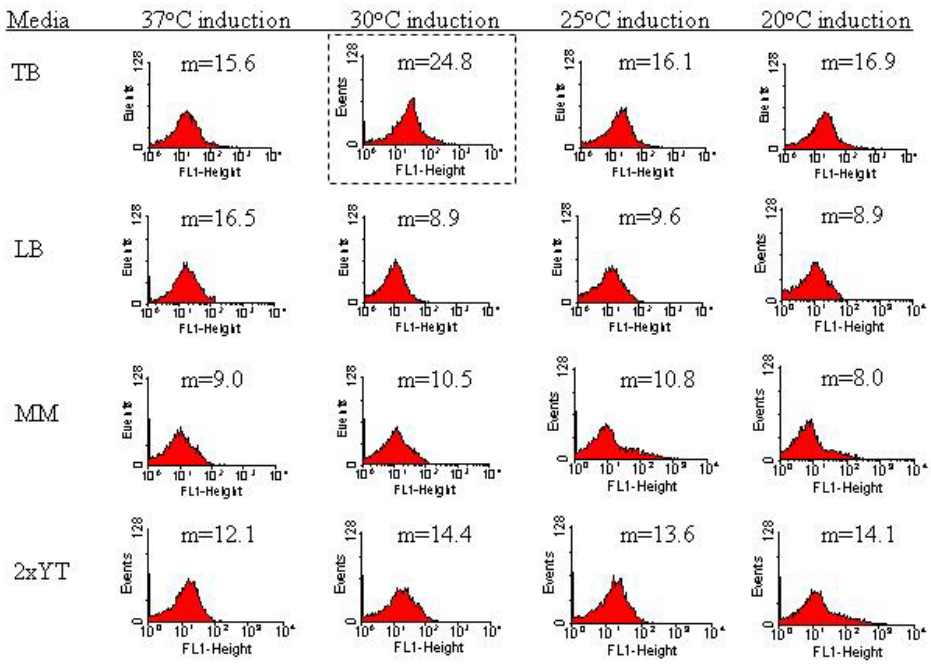

Figure 3.3. Media and temperature effects on cell fluorescence. *E. coli* FA113 cells containing the mutant C8 ssTorA-26-10 construct were grown at 37°C for one hour, protein synthesis was induced by arabinose and the cells were transferred to a different temperature as indicated and incubated for an additional four hours. The fluorescence histogram for this control condition, i.e. the conditions used in other studies reported in this thesis, is shown in the box. Each column represents a different four hour induction temperature used after growth at 37°C for one hour. Each row represents the media used for those cultures. Each culture was run in triplicate, and the histograms provided are representative of each triplicate set of samples. m, geometric mean fluorescence intensity of the histogram.

The flow cytometric results are shown in Figure 3.3. Not only were the highest cell densities achieved in TB media, but also the highest fluorescence was observed under the conditions used previously in this thesis. TB media is the most nutrient rich media tested, thus allowing the cells to reach higher densities. Using TB media as the starting point, we proposed that the switch from growth at 37°C to induction at 30°C might cause a stress response in the cells. Therefore, cultures were kept at a constant temperature for both their growth and induction phases. Adjusting the media pH was also examined since the growth pH can affect the PMF that is maintained by the cell by altering the

concentration differences of the ions on each side of the inner cell membrane. Since Tat export relies on the PMF for the energy necessary for translocation, a cell that can establish a higher PMF may have increased export.

As shown in Figure 3.4, growth at 37°C in TB media at pH 7.2 followed by protein synthesis at 30°C produced the highest fluorescence. Growth and induction at a lower temperature resulted in a lower level of fluorescence that was observable in the triplicate samples for each conditions. Growth at a more basic or acidic pH also caused a decrease in the fluorescence of the cells. A molecular basis for the effect of pH on export could not be proposed since the growth pH exerts a very complex effect on cell physiology.

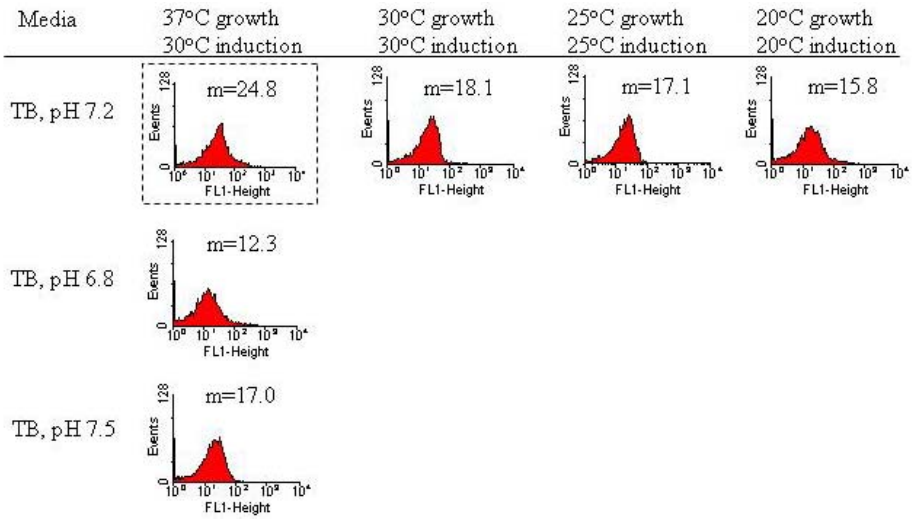

Figure 3.4. Temperature and pH effects on cell fluorescence. *E. coli* FA113 cells transformed with pTorA-2610$_{C8}$ were grown in TB media at the conditions indicated. Following growth at the indicated temperature, protein synthesis was induced with arabinose, and the cells were incubated at the indicated temperature for an additional four hours. The fluorescence histogram for the optimal growth conditions is shown in the box. Each column represents a different one-hour growth and four hour induction temperature. Each fluorescence histogram is a representative sample from a set of triplicate cultures. m, geometric mean fluorescence intensity of the histogram.

3.3.3 Fluorescence in various oxidizing strains

Since scFv proteins have two intramolecular disulfide bonds, an *E. coli* strain that has an oxidizing cytoplasm is used to allow oxidative protein to form [1, 24, 25]. As discussed in Chapter 2, strains with an oxidizing cytoplasm are necessary in order for the scFv protein to be competent for export through the Tat pathway. In *E. coli*, the *trx* and *gor* genes play a key role in maintaining the reducing nature of the cytoplasm [45]. Thioredoxin reductase reduces thioredoxins which in turn reduce disulfide bonds [46, 47]. Similarly, glutathione is reduced by the glutathione reductase enzyme encoded by the *gor* gene [48].

However, maintaining an oxidizing cytoplasm adversely affects the growth rate and cell viability of the FA113 strain since native proteins that are normally reduced are now partially oxidized, affecting their function. Recently the Beckwith laboratory created a set of strains having different mutations that result in an oxidizing cytoplasm. Different *gor* and *trx* genes have deletions in the genes encoding different components of the thioredoxin or glutathione reduction pathways. Specifically, strain MJF256.10 is lacking the *gshA* gene that codes for gamma-glutamate-cysteine ligase, the enzyme involved in glutathione synthesis [24]. Without glutathione, glutathione reductase has no substrate to reduce thereby inactivating the glutathione reduction pathway. The *trxB* and *gor* genes have been deleted from the chromosome in strain SMG96 rather than having mutations in them as in FA113. Strain DR600 is similar except on the *trxB* gene had been deleted from the chromosome. Strains DR602 and DR611 are deficient in the *trxA* and *trxC* genes along with *trxB* and *gor* genes present in the FA113 strain. These two additional genes encode for thioredoxin and thioredoxin2, respectively, and their deletion prevents the accumulation of oxidized thioredoxins in the cytoplasm.

The pTorA-2610$_{C8}$ plasmid was used to transform *E. coli* strains DR600, DR602, DR611, MJF256.10 and SMG96. Transformants were grown in TB media and induced as above and flow cytometry was used to analyze cell fluorescence. As seen in Figure 3.5, *E. coli* FA113 had the highest fluorescence intensity followed closely by strains DR611, DR602 and DR600. The optical densities of the cell cultures of the tested strains were all less than those attained by *E. coli* FA113. *E. coli* FA113 achieved a final optical density of 4.1 at 600 nm while *E. coli* strains DR600, DR602, DR611, SMG96 and MJF256.10 achieved densities of 2.2, 0.3, 0.4, 1.5 and 2.2 at 600 nm, respectively. Therefore, although their fluorescent intensities were similar to that of FA113, their lower optical densities of the liquid cell cultures translates to a significantly lower overall yield of scFv protein.

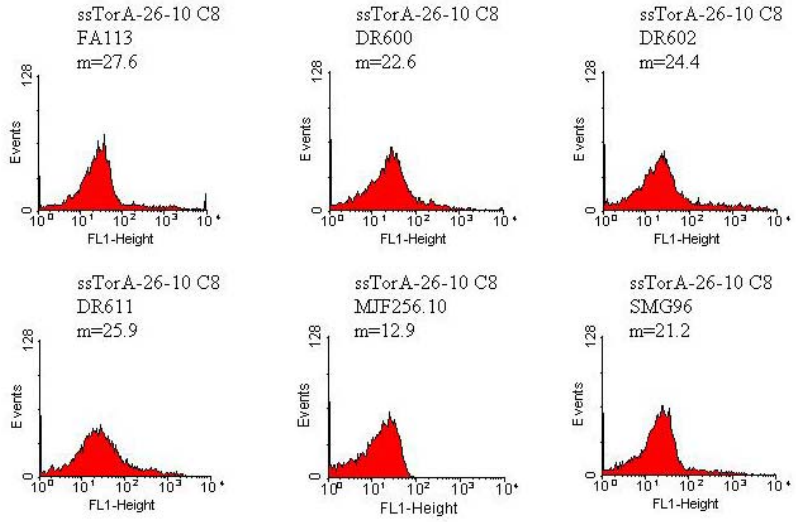

Figure 3.5. Cell fluorescence in different *E. coli* strains containing mutations that render the cytoplasm oxidizing. Plasmid pTorA-2610$_{C8}$ that encodes the mutant C8 scFv was used to transform various *E. coli* oxidizing cytoplasm strains. *E. coli* FA113 which was used in the other studies described here is shown as a control. m, geometric mean fluorescence intensity of the histogram.

3.4 DISCUSSION

In this study, we sought to explore methods to increase the efficiency of export through the Tat pathway to render it a more useful means of protein production. The low yields of exported protein have hampered the use of Tat in biotechnology even though it offers a quality control advantage in that it exports only fully and correctly folded proteins to the periplasm. Using the 26-10 and C8 scFvs as reporters, we first evaluated the effect of co-expression of several proteins known or suspected to affect Tat export. Cell fluorescence, due to the binding of a fluorescent hapten by the scFv antibodies in the periplasm, was used to assay the amount of Tat secreted protein. The largest increase in protein export was observed by co-expressing the three proteins that comprise the Tat translocon, namely TatA, TatB and TatC. A 2.4-fold increase in fluorescence was observed indicating that the rate limiting step in export may be the number of Tat pores available for protein export. By providing more translocon complexes, more competent scFv protein molecules can interact with the complex and be exported. In addition, we also investigated the effect of temperature, pH and growth media on the growth and fluorescence intensity of cells expressing the C8 ssTorA-26-10 scFv. However, we found that growth in TB media of cells incubated at 37^{o}C and then shifted to 30^{o}C for induction of protein synthesis was optimal for both achieving the highest cell density and maximum fluorescence, and hence, protein export per cell. Although lower induction temperatures could allow the scFv to be more soluble, we found that the fluorescence intensity decreased. This could indicate that the lower temperatures may slow the rate of expression and negate any beneficial effects gained by increased solubility.

Finally, we considered the use of alternate strains of *E. coli* which contained mutations in different *gor* and *trx* genes encoding different components of the thioredoxin or glutathione reduction pathways which normally maintain the native, reducing

cytoplasm. Essential proteins and enzymes may be oxidized when these pathways are inactivated leading to a reduction in cell growth and viability. These strains were recently created to supply the oxidizing cytoplasm necessary for disulfide bond formation while attempting to reduce the negative side effects of the oxidizing cytoplasm. These strains were transformed with the pTorA-2610$_{C8}$ plasmid, grown and induced as described above, and analyzed by flow cytometry. Although their cell fluorescent intensities were similar to that of strain FA113, their liquid cell cultures achieved lower optical densities which translates to a significantly lower overall yield of scFv protein.

As a benchmark, a 30-fold increase in fluorescence is seen over background fluorescence when 26-10 is exported via the Sec pathway [26]. The current maximum fluorescence achieved in this study was 16-fold over background. Hence, other methods of improvement need to be investigated in order to increase the level of Tat export to that of the Sec pathway. One of these methods may be to engineer a better strain based on the *E. coli* FA113 strain. For example, an additional TatABC operon under the control of an inducible promoter could be placed on the chromosome. Therefore, an additional protein like TorD could be co-expressed with the C8 ssTorA-26-10 fusion while TatABC is overexpressed from the new operon. This approach could yield increased fluorescence levels comparable to the those obtain via Sec export.

3.5 REFERENCES

1. DeLisa, M.P., D. Tullman, and G. Georgiou, *Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6115-20.
2. Thomas, J.D., et al., *Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli.* Mol Microbiol, 2001. **39**(1): p. 47-53.
3. Barth, S., et al., *Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions.* Appl Environ Microbiol, 2000. **66**(4): p. 1572-9.
4. Cupit, P.M., et al., *Cloning and expression of single chain antibody fragments in Escherichia coli and Pichia pastoris.* Lett Appl Microbiol, 1999. **29**(5): p. 273-7.
5. Yahr, T.L. and W.T. Wickner, *Functional reconstitution of bacterial Tat translocation in vitro.* Embo J, 2001. **20**(10): p. 2472-9.
6. DeLisa, M.P., et al., *Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway.* J Bacteriol, 2004. **186**(2): p. 366-73.
7. Chanal, A., C.L. Santini, and L.F. Wu, *Specific inhibition of the translocation of a subset of Escherichia coli TAT substrates by the TorA signal peptide.* J Mol Biol, 2003. **327**(3): p. 563-70.
8. Mikhaleva, N.I., et al., *Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in Escherichia coli.* FEBS Lett, 1999. **463**(3): p. 331-5.
9. Li, S.Y., B.Y. Chang, and S.C. Lin, *Coexpression of TorD enhances the transport of GFP via the TAT pathway.* J Biotechnol, 2006. **122**(4): p. 412-21.
10. De Keersmaeker, S., et al., *Evaluation of TatABC overproduction on Tat- and Sec-dependent protein secretion in Streptomyces lividans.* Arch Microbiol, 2006.
11. Loll, P.J., *Membrane protein structural biology: the high throughput challenge.* J Struct Biol, 2003. **142**(1): p. 144-53.
12. Gohlke, U., et al., *The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter.* Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10482-6.
13. Lee, P.A., D. Tullman-Ercek, and G. Georgiou, *The Bacterial Twin-Arginine Translocation Pathway.* Annu Rev Microbiol, 2006.
14. Segatori, L., et al., *Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in Escherichia coli: Reconciling two competing pathways.* Proc Natl Acad Sci U S A, 2004. **101**(27): p. 10018-23.
15. Genest, O., et al., *Signal peptide protection by specific chaperone.* Biochem Biophys Res Commun, 2006. **339**(3): p. 991-5.
16. Schwan, W.R., et al., *Osmolarity and pH growth conditions regulate fim gene transcription and type 1 pilus expression in uropathogenic Escherichia coli.* Infect Immun, 2002. **70**(3): p. 1391-402.

17. King, P.W. and A.E. Przybyla, *Response of hya expression to external pH in Escherichia coli.* J Bacteriol, 1999. **181**(17): p. 5250-6.
18. Wyborn, N.R., et al., *Expression of the Escherichia coli yfiD gene responds to intracellular pH and reduces the accumulation of acidic metabolic end products.* Microbiology, 2002. **148**(Pt 4): p. 1015-26.
19. Bailey, A.M., M.A. Webber, and L.J. Piddock, *Medium plays a role in determining expression of acrB, marA, and soxS in Escherichia coli.* Antimicrob Agents Chemother, 2006. **50**(3): p. 1071-4.
20. Herendeen, S.L., R.A. VanBogelen, and F.C. Neidhardt, *Levels of major proteins of Escherichia coli during growth at different temperatures.* J Bacteriol, 1979. **139**(1): p. 185-94.
21. Gough, K.C., et al., *Selection of phage antibodies to surface epitopes of Phytophthora infestans.* J Immunol Methods, 1999. **228**(1-2): p. 97-108.
22. Baneyx, F., *"In vivo Folding of Recombinant proteins in Escherichia coli" in Manual of Industrial Microbiology and Biotechnology*. 2 ed, ed. D.A. Davies JE, Cohen G, Hershberger CL, Forney and H.I. LJ, Hu W-S, Wu J-HD, Sherman DH, Wilson RC. 1999: ASM Press.
23. Meyer-Rosberg, K., et al., *The effect of environmental pH on the proton motive force of Helicobacter pylori.* Gastroenterology, 1996. **111**(4): p. 886-900.
24. Bessette, P.H., et al., *Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13703-8.
25. Martineau, P. and J.M. Betton, *In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm.* J Mol Biol, 1999. **292**(4): p. 921-9.
26. Chen, G., et al., *Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS).* Nat Biotechnol, 2001. **19**(6): p. 537-42.
27. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.
28. Allen, S.C., et al., *Essential cytoplasmic domains in the Escherichia coli TatC protein.* J Biol Chem, 2002. **277**(12): p. 10362-6.
29. Brissette, J.L., et al., *Phage shock protein, a stress protein of Escherichia coli.* Proc Natl Acad Sci U S A, 1990. **87**(3): p. 862-6.
30. Kleerebezem, M., W. Crielaard, and J. Tommassen, *Involvement of stress protein PspA (phage shock protein A) of Escherichia coli in maintenance of the protonmotive force under stress conditions.* Embo J, 1996. **15**(1): p. 162-71.
31. Model, P., G. Jovanovic, and J. Dworkin, *The Escherichia coli phage-shock-protein (psp) operon.* Mol Microbiol, 1997. **24**(2): p. 255-61.
32. Kleerebezem, M. and J. Tommassen, *Expression of the pspA gene stimulates efficient protein export in Escherichia coli.* Mol Microbiol, 1993. **7**(6): p. 947-56.
33. De Lorenzo, F., et al., *Purification and properties of an enzyme from beef liver which catalyzes sulfhydryl-disulfide interchange in proteins.* J Biol Chem, 1966. **241**(7): p. 1562-7.

34. Chen, J., et al., *Chaperone activity of DsbC.* J Biol Chem, 1999. **274**(28): p. 19601-5.
35. Kurokawa, Y., H. Yanagi, and T. Yura, *Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in Escherichia coli.* J Biol Chem, 2001. **276**(17): p. 14393-9.
36. Qiu, J., J.R. Swartz, and G. Georgiou, *Expression of active human tissue-type plasminogen activator in Escherichia coli.* Appl Environ Microbiol, 1998. **64**(12): p. 4891-6.
37. Zhang, Z., et al., *Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in Escherichia coli.* Protein Expr Purif, 2002. **26**(2): p. 218-28.
38. Pommier, J., et al., *TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in Escherichia coli.* J Biol Chem, 1998. **273**(26): p. 16615-20.
39. Jack, R.L., et al., *Coordinating assembly and export of complex bacterial proteins.* Embo J, 2004. **23**(20): p. 3962-72.
40. Jack, R.L., et al., *Constitutive expression of Escherichia coli tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth.* J Bacteriol, 2001. **183**(5): p. 1801-4.
41. Mori, H. and K. Cline, *A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase.* J Cell Biol, 2002. **157**(2): p. 205-10.
42. Berks, B.C., T. Palmer, and F. Sargent, *The Tat protein translocation pathway and its role in microbial physiology.* Adv Microb Physiol, 2003. **47**: p. 187-254.
43. Barrett, C.M., et al., *Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in Escherichia coli.* Biochem Biophys Res Commun, 2003. **304**(2): p. 279-84.
44. Martinez, M.B., M.C. Flickinger, and G.L. Nelsestuen, *Steady-state enzyme kinetics in the Escherichia coli periplasm: a model of a whole cell biocatalyst.* J Biotechnol, 1999. **71**(1-3): p. 59-66.
45. Carmel-Harel, O. and G. Storz, *Roles of the glutathione- and thioredoxin-dependent reduction systems in the Escherichia coli and saccharomyces cerevisiae responses to oxidative stress.* Annu Rev Microbiol, 2000. **54**: p. 439-61.
46. Holmgren, A., et al., *Thiol redox control via thioredoxin and glutaredoxin systems.* Biochem Soc Trans, 2005. **33**(Pt 6): p. 1375-7.
47. Masip, L., et al., *An engineered pathway for the formation of protein disulfide bonds.* Science, 2004. **303**(5661): p. 1185-9.
48. Fernandes, A.P. and A. Holmgren, *Glutaredoxins: glutathione-dependent redox enzymes with functions far beyond a simple thioredoxin backup system.* Antioxid Redox Signal, 2004. **6**(1): p. 63-74.

Chapter 4

FACS Screening for the Selection of Mutant Proteins Resistant to Degradation by the AAA+ Protease ClpXP

4.1 INTRODUCTION

As described in the first chapter, proteolysis is used to regulate response to environmental factors and for the disposal of proteins that have become unfolded and/or chemically damaged [1-4]. ATP-dependent proteases are present in most cells and subcellular compartments such as the eukaryotic proteasome, Clp proteases in bacteria, Lon proteases in mitochondria and bacteria and membrane-bound proteases in mitochondria (Yta10/12 and Yme1), chloroplasts (FtsH) and bacteria (FtsH or HflB) [5]. The Clp proteases in bacteria ClpAP, ClpXP and ClpYQ are responsible for proteolysis of sigma factors that play a role in regulating gene expression [6] and degradation of damaged proteins and proteins from ribosomes that have stalled during translation [7]. The serine-type protease ClpP of *E. coli* is the best studied of these prokaryotic proteases. It is composed of two seven-subunit rings comprised of 14 identical subunits with an internal cavity only accessible through narrow axial channels [8]. Since this channel is only accessible to small peptides, the proteolytic capability of ClpP is limited [9].

ClpP associates with the ATP-dependent chaperones, ClpA or ClpX. The ClpAP or ClpXP complex can degrade folded proteins. ClpA and ClpX are composed of six subunits arranged in a hexagonal ring that can associate with either or both sides of ClpP, but recognize different degradation tags [10, 11]. Both have the role of binding the target protein by recognizing a degradation tag, unfolding the protein, and translocating the

protein into the proteolytic chamber of ClpP [12, 13]. Studies by Lee *et. al.* [5] used mutant proteins in which the C-terminus was connected to the polypeptide chain by a linker at various points to create a circular protein. Degradation tags were then attached to the N-terminus and also the C-terminus for all of the mutants. Degradation of the circular proteins did not depend on the global stability of the protein, as determined by the free energy of unfolding, but rather by the local structure next to the degradation tag. The conclusion of this research was that the Clp system degrades proteins by progressively unraveling them from their degradation signal. Also, Burton *et al.* [14] determined that, even for a wide range of protein stabilities, the ClpXP protease degraded each mutant at a rate that varied by less than a factor of two. This indicated that proteins are unfolded by the application of a mechanical force to the degradation tag causing the unfolding of the protein structure neighboring the tag [14, 15].

We hypothesized that since the ClpA and ClpX chaperones unravel proteins by applying a mechanical force to unfold them, the ClpXP system could be utilized as a method to isolate proteins that are more resistant to shear. In other words, proteins that are more resistant to degradation by the ClpXP system are therefore more resistant to the shear force used to unfold the protein. Proteins that can withstand higher shear forces before they unfold can be useful in applications where anchored proteins are subjected to blood flow [16] and cell adhesion by receptor-ligand interactions that are also subjected to shear forces in blood flow [17]. For example, research is being conducted into methods to increase the adhesion of endothelial cells to surfaces by changing the binding properties of intergrins, a class of cell adhesion molecules that bind to ligands containing the arginine-glycine-aspartic acid (RGD) peptide sequence [18]. In addition, these intergrins are also used to bind different proteins involved in wound healing [19]. In both cases, the cells or proteins are subject to physiologically relevant shear forces induced by

blood flow. An improved protein that does not unfold under these conditions so that it maintains its function can lead to better attachment of cells to biomedical implants or improved wound-healing treatments. Researchers are currently mutating single amino acids near the RGD binding site in order to increase the binding of these ligands. However, mutations in the framework regions surrounding the binding sites may also increase binding affinities by stabilizing the protein's conformation and therefore the binding site, as is the case with antibodies [20, 21]. Therefore, a screen that isolates mutants more resistant to mechanical unfolding would be beneficial in order to obtain proteins that can bind ligands under the shear conditions associated with blood flow.

In addition to engineering more stable proteins, being able to control the Clp system in *E. coli* may allow greater yields of protein since cytoplasmic proteolysis would be reduced [22]. For example, recombinant proteins that are overexpressed in *E. coli* may aggregate and induce degradation by the Clp system. Inhibiting the Clp system at the start of protein induction would ensure that recombinant protein yields would not be decreased by degradation. Similarly, eukaryotic cells also contain proteases such as the proteasome that degrades abnormal polypeptides [23]. Therefore, study of the Clp system and its control may lead to a similar control of the proteasome which then could be applied to increasing protein production in eukaryotic cells.

In this study, we employed a genetic strategy that capitalizes on FACS screening and was aimed at: (i) identifying genes that, upon multicopy expression, could inhibit degradation by the Clp machinery and (ii) isolating intragenic mutations that conferred higher resistance of a model protein to degradation by the Clp machinery. For the first study, a genomic library was constructed and screened using a fluorescent reporter system discussed below. Several clones exhibiting higher fluorescence were isolated, however, further experiments did not yield any regulators or inhibitors of the Clp system.

In addition, several libraries of GFP were constructed and targeted to the Clp system using different degradation signals. It was proposed that different mutations may increase the stability of GFP to degradation by Clp thus conferring higher cell fluorescence. Since mechanical stability is an important parameter in the degradation of proteins by the Clp machinery, mutants exhibiting higher resistance to Clp would be likely to have improved mechanical properties. This could, in turn, provide a system for engineering proteins that are more resistant to shear force. However, after several attempts, no mutant of GFP was isolated that showed any increase in fluorescence when used in the same reporter system. This indicated that no mutations could stabilize GFP, while maintaining its function, against the Clp system.

4.2 MATERIALS AND METHODS

4.2.1 Strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table 4.1. *E. coli* XL1-Blue cells were used for plasmid propagation.

To construct the pSsrA-GFP pBAD33 plasmid, primers SsrA.f and SsrA.r were used to amplify the *SsrA* tag from pGFP-SsrA by PCR. The PCR fragment and pGFP-SsrA were digested by *Kpn*I and *Sal*I and ligated forming a pBAD33 *ssrA-ssrA* construct. This new construct was digested with *Sal*I and *Sph*I to remove the second *ssrA* tag. PCR was performed with primers GFP.f and GFP.r and pGFP-SsrA as template to yield fragments of full length *gfp* that were digested with *Sal*I and *Sph*I and ligated into the already cut *ssrA-ssrA* plasmid to create pSsrA-GFP.

Using primers Xho.f, Xho.r, Sapa48.f, and Sapa49.r and pGFP-SsrA as template, overlap PCR was carried out to introduce an *Xho*I restriction site at position 557 of 714 of *gfp* by changing the WT sequence from *CTCCAA* to *CTCGAG*. The fragment containing the full length *gfp* with the *Xho*I restriction site was digested by *Kpn*I and *Sal*I and ligated into similarly digested pGFP-SsrA to yield $pGFP_{xho}$-SsrA.

pGFP-RepA was constructed by PCR with primers Xho.f and RepA.r and pGFPxho-SsrA as template to yield a fragment containing the end of *gfp* with a *RepA* tag. This fragment and $pGFP_{xho}$-SsrA were digested with *Xho*I and *Sph*I, and upon ligation, pGFP-RepA was created.

Table 4.1. Strains, plasmids, and primers used in this study

E. coli Strain or Plasmid	Relevant Genotype or Features	Source
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI*^{q} *ZΔM15* Tn*10* (Tet^r)]	Lab collection
pBAD33	*araBAD* promoter, pBR322 *ori*, cm^r	Ref. [24]
pTrc99A	*trc* promoter, ColE1 *ori*, cm^r	Amersham Biosciences
pGFP-SsrA	GFP-SsrA in pBAD33	Lab collection
pSsrA-GFP	SsrA-GFP in pBAD33	This study
pGFP$_{xho}$-SsrA	GFP-SsrA in pBAD33. GFP has engineered internal *Xho*I restriction site.	This study
pGFP-RepA	GFP-RepA in pBAD33	This study
pEno	Eno in pTrc99A	This study
pPyrG	PyrG in pTrc99A	This study
pPspA	PspA in pTrc99A	Lab collection
pPspB	PspB in pTrc99A	This study
pPckA	PckA in pTrc99A	This study
pSgcX	SgcX in pTrc99A	This study
Primers		
Eno.f	CCAGCCATGGATGTCCAAAATCGTAAAAATCATC	
Eno.r	CCAGAAGCTTTTATGCCTGGCCTTTGAT	
PyrG.f	CCAGCCATGGATGACAACGAACTATATTTTTGTG	
PyrG.r	CCAGAAGCTTTTACTTCGCCTGACG	
PspB.f	CCAGCCATGGATGAGCGCGCTATTT	
PspB.r	CCAGAAGCTTTTAGCGATCCCTCCA	
PckA.f	CCAGCCATGGATGCGCGTTAACAATGGT	
PckA.r	CCAGAAGCTTTTACAGTTTCGGACCAGC	
SgcX.f	CCAGCCATGGGTGATTGTTTTACGGTCTACG	
SgcX.r	CCAGAAGCTTTCATAATGGATGTGCCTCTTG	
Sapa48.f	AGAGGAGAAAGGTACCCATG	
Sapa49.r	CGTCGTTTGCTGCGTCGAC	
Sapa50.r	CGTCGTTTGCTGCGTCGACTTTGTA	
SsrA.f	CTTTGGTACCATGGCAGCAAACGACGAAAAC	
SsrA.r	GACGGTCGACAGCTGCTAAAGCGTAGTT	
GFP.f	ATGTTTGTCGACAGTAAAGGAGAAGAACTTTTCACT	
GFP.r	TTCGAAGCATGCTTATTTGTATAGTTCATCCATGCCAT	
Xho.f	CAAAATACTCGAGTTGGCGATGG	
Xho.r	CATCGCCAACTCGAGTATTTTGTT	
RepA.r	TTCGAAGCATGCTTATTCAATGTCTGCGTAAAGAATATCGGAGATAAATGAT TGATTCATGTCGACTTTGTATAGTTCATCCAT	

Constructs pEno, pPyrG, pPspB, pPckA, and pSgcX were made by cloning each gene from XL1-Blue genomic DNA using each gene's forward and reverse primer as listed in Table 4.1. Empty pTrc99A plasmid and each cloned gene was digested by *Nco*I and *Hind*III and ligated to create each new plasmid.

4.2.2 Library Construction

A genomic library was constructed from DNA isolated from XL1-Blue cells using a Qiagen DNAeasy Kit. The DNA was digested with *Sau3*I for 30 minutes which was determined to yield fragments approximately 1,000 bases in length. Empty pTrc99A plasmid was digested with *BamH*I and then treated with calf intestinal alkaline phosphatase (CIAP) to prevent self-ligation of the plasmid. The genomic DNA fragments were ligated into the pTrc99A backbone, and the completed ligation mixture was electroporated into XL1-Blue cells already containing pGFP-SsrA. $1x10^6$ transformants were obtained.

Mutant *gfp-ssrA* library I was constructed using error-prone PCR [25] with primers Sapa48.f and Sapa49.r and pGFP-SsrA as template. After digestion with *Kpn*I and *Sal*I, the resulting DNA fragments and pGFP-SsrA were ligated and electroporated in XL1-Blue. Five clones were sequenced and averaged six mutations per gene for an error rate of 0.7%. The resulting library size was $2x10^6$ transformants.

Mutant *gfp-ssrA* library II was constructed in the same manner except that Sapa50.r was used to replace Sapa49.r to prevent the mutation of either of the last two codons to become stop codons. Nine clones were sequenced from the resulting $5.1x10^6$ library, averaging 2 mutations per *gfp* gene or a 0.3% error rate.

The final two *gfp-ssrA* libraries, libraries III and IV, were constructed using error-prone PCR [25] with Sapa48.f and Xho.r primers with $pGFP_{xho}$-SsrA as template. This allowed only mutations to occur in the first 557 of 714 nucleotides of *gfp*. The PCR products and $pGFP_{xho}$-SsrA were digested with *Kpn*I and *Xho*I, ligated, and electroporated into XL1-Blue. Ten clones were sequenced from the resulting $2x10^7$ library III for an average of 8 changes per *gfp* or an error rate of 1.4%. The size of library

IV was $5x10^6$ transformants and was created with a lower error rate of 0.3% or 1.7 mutations per gene as determined from the sequencing of nine clones selected at random.

4.2.3 Growth conditions, expression and flow cytometry

Unless otherwise noted, cells were grown at 37°C on LB media with 25 μg/ml chloramphenicol and/or 100 μg/ml ampicillin, as appropriate.

For sorting of the genomic library, 100 ml of LB media containing 25 μg/ml chloramphenicol and 100 μg/ml ampicillian was inoculated with 1 ml of library frozen stock. After growth with shaking for 1 hour at 37°C, protein expression was induced by the addition of arabinose and isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.2% and 0.1 mM, respectively. Following a 5 hour induction period, 1 ml of culture was collected, pelleted at 2,000 g for 5 minutes, and resuspended in 1 ml of 1x phosphate buffered saline (PBS) before sorting. Sorting of the mutant GFP libraries I-IV was carried out in the same manner except only chloramphenicol was used in the LB media and arabinose as the inducer. Sorting was carried out on a Becton-Dickinson FACSort flow cytometer with settings as described in Chapter 2 with the exception that distilled water was run as sheath for all scans.

Western blotting was performed as previously described [26] with slight changes. After induction, 1 ml of cells were pelleted at 2,000 g for 5 minutes, washed in 1xPBS, and pelleted again. The pellet was then mixed with 250 μl 2xSDS loading buffer containing 2-mercaptoethanol. After boiling for 10 minutes, the samples were loaded and run on a 12% Tris-glycine precast gel (Cambrex). Sample volumes were adjusted to represent equivalent cell numbers as determined by optical density measurements of each culture at the time of collection. Proteins were then transferred to a polyvinylidine fluoride membrane using a semi-dry transfer graphite electroblotting system (Millipore). The membranes were then blocked for 1 hour at room temperature or overnight at 4°C in

5% nonfat dry milk/TBST buffer (15 mM Tris-HCl, 4 mM Tris, 137 mM NaCl, 0.1% Tween-20, pH 7.6). Antibodies were diluted in 1% nonfat dry milk/TBST for 1 hour at room temperature. The primary antibody used for Western blots was a monoclonal mouse anti-GFP (Clonetech).

For flow cytometric analysis, cells containing the appropriate constructs were grown overnight in LB media with the appropriate antibiotics. 250 µl of overnight culture was used to inoculate 10 ml of fresh media, and growth and induction conditions were the same as those described above. Cells were harvested from 1 ml of culture by centrifugation at 2,000 g and resuspended in 1 ml of 1xPBS. Depending upon cell density, 3 to 10 µl of each sample was added to 1 ml of 1xPBS in a sieve tube before being analyzed on a Becton-Dickinson FACSort flow cytometer with the above settings. The fluorescence of individual clones was also analyzed on a fluorescence plate reader. Colonies were picked to individual wells of a 96-well plate (Costar brand round bottom, polystyrene, non-binding surface) containing 100 µl of LB media with the appropriate antibiotics. After overnight growth at 37^{o}C, new plates were subcultured by replica plating, grown for 1.5 hours at 37^{o}C, and induced by the addition of arabinose and/or IPTG to a final concentration of 0.2% and 0.1 mM, respectively. Following the induction period, the plates were centrifuged at 2,000 g for ten minutes at room temperature. Cells in each well were resuspended in 100 µl of 1xPBS before being analyzed by a BioTek Synergy™ HT Microplate Reader using an excitation filter of 480 nm and an emission filter of 530 nm. To compare fluorescence readings measured on microtiter plates with the fluorescence distribution of cells measured by flow cytometry, 5 µl of sample from a well was added to 1 ml of 1xPBS in a sieve tube to be analyzed by flow cytometry.

4.3 RESULTS

4.3.1 Fluorescent reporter system to monitor degradation by ClpXP

GFP has been used extensively as a fluorescent reporter in *E. coli* [27-29], is easily overexpressed and folds correctly in the cytoplasm [30, 31]. By adding the C-terminal SsrA peptide tag (AANDENYALAA) to GFP, the GFP-SsrA fusion is targeted to degradation by the cytoplasmic protease ClpXP [32, 33] (see Figure 4.1, adapted from [34]). Due to the degradation of GFP-SsrA by ClpXP, the fluorescence of cells expressing GFP is much greater than cells expressing the GFP-SsrA fusion providing a high signal-to-noise ratio as shown in Figure 4.1B.

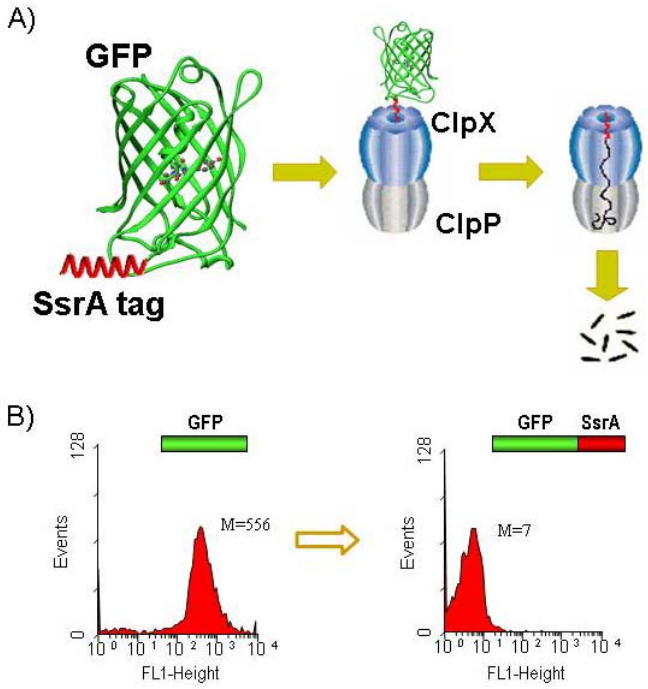

Figure 4.1. Fluorescent reporter system. (A) GFP expressed with an SsrA tag is targeted and degraded by ClpXP. (B) Sample histograms showing fluorescence of GFP and GFP-SsrA expressed from a plasmid. As indicated, due to the SsrA tag, the GFP is degraded, lowering the fluorescent signal.

The GFP-SsrA fusion was expressed under the control of the pBAD promoter in pBAD33 plasmid [24]. Expression of the fusion protein occurs upon induction of the cells with the addition of 20% arabinose. However, high-level expression of the fusion may saturate the ClpXP machinery which is only synthesized at constitutive levels. Thus, the time-course of GFP-SsrA in *E. coli* XL1-Blue cells transformed with pGFP-SsrA was determined by inducing cultures in mid-log phase with arabinose and analyzing the cell fluorescence in samples taken hourly by flow cytometry. As shown in Figure 4.2, approximately five hours after induction, most of the GFP-SsrA fusion protein had been degraded, and the fluorescence had returned to background levels. Therefore, five hours after the onset of GFP-SsrA synthesis, sorting by FACS for higher fluorescent cells than background could be expected to yield clones of interest.

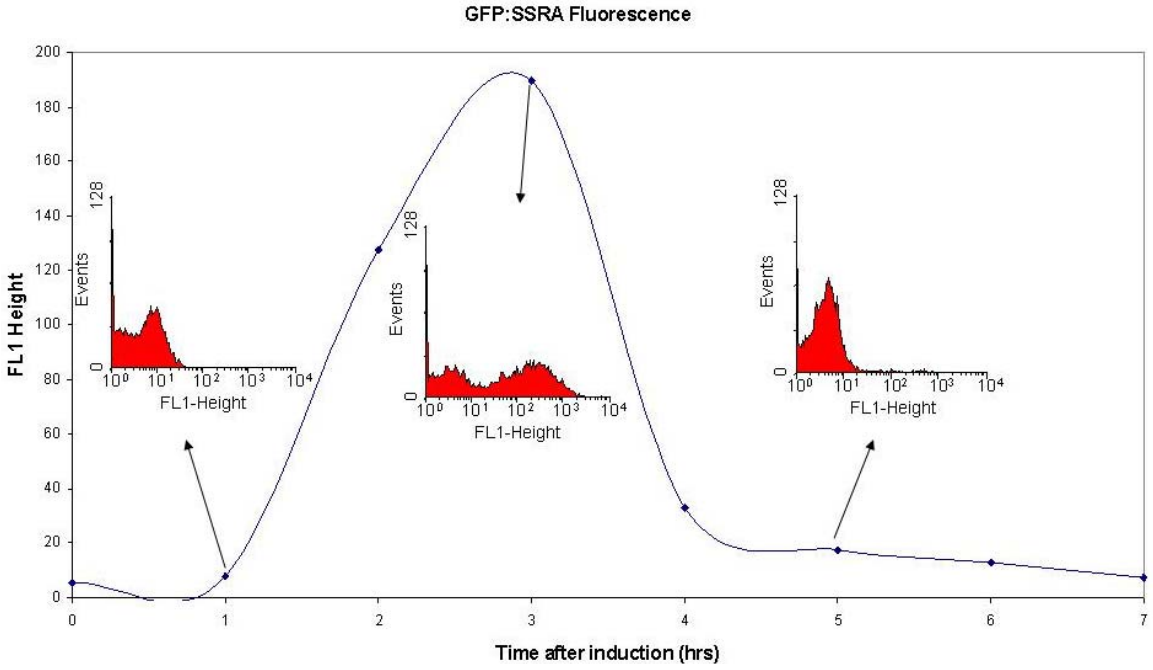

Figure 4.2. Time dependent fluorescence of GFP-SsrA. The synthesis of GFP-SsrA was induced at t=0 and samples were taken hourly and analyzed by flow cytometry. Sample histograms showing the fluorescence of the cell populations are shown at hours 1, 3 and 5. A smoothed line curve connects the data points.

4.3.2 Genomic library screening for regulators of ClpXP

We sought to determine if any genes present in the *E. coli* genome act as repressors of the Clp operon or inhibitors of the Clp machinery itself. A genomic library was constructed and co-expressed with GFP-SsrA. In a genomic library, the entire genome of an organism is broken into smaller DNA fragments which are ligated into a vector, typically a plasmid. Gene fragments resulting in higher GFP-SsrA fluorescence may encode for factors that regulate or inhibit the Clp machinery.

A genomic library was created by digesting *E. coli* genomic DNA followed by ligation into pTrc99A. Digestion times were adjusted to yield DNA fragments of approximately 1,000 base pairs to ensure each fragment contains a single, intact gene or the majority of one gene. The ligation mixture was transformed into *E. coli* XL1-Blue cells already containing pGFP-SsrA, and resulted in a library of 10^6 transformants. Since the *E. coli* genome has 4,405 genes [35], each gene should be statistically represented at this library size.

The synthesis of GFP-SsrA and any polypeptides encoded by the library gene fragment were induced and, at five hours later, the library was sorted by FACS. 1.4×10^7 cells were sorted, and the top 1% of fluorescent events consisting of 1,275 were collected. The collected cells were grown overnight, and 500 colonies were then transferred to a 96-well plate. After subculturing to another 96-well plate for growth for one hour and induction for five hours, cell fluorescence was determined on a fluorescence plate reader. Ten clones with fluorescence higher than the WT GFP-SsrA alone were selected, grown overnight, and purified plasmid was obtained from each culture. These purified plasmids were used to back-transform new *E. coli* XL1-Blue cultures. Back-transformation ensures that the observed phenotypes are due to the expression of the

cloned gene and not from other effects which may have arisen during the growth of the culture.

Induction of protein synthesis and analysis by flow cytometry revealed that the isolated clones were highly fluorescent as compared to the fluorescence of cells transformed with pGFP-SsrA together with empty Trc99A plasmid and grown under identical conditions (see Figure 4.3). The genomic insert DNA in the pTrc99A plasmid was sequenced, and the sequences were used in a BLAST search [36]. Matches to the following genes were found: *eno*, an enolase used in energy metabolism; *pyrG*, a CTP synthetase; *pspA*, an inner membrane phage shock protein; *pspB*, an activator of pspA; *pckA*, a phosphoenolpyruvate carboxykinase; and *sgcX*, a putative lyase/synthase (all gene definitions from [37]]). The remaining four sequences were for ribosomal proteins as overexpression of ribosomal proteins could increase translation [38] and overwhelm the ClpXP machinery. Therefore, more GFP-SsrA would be present at the end of the five hour induction period, leading to higher fluorescence.

However, only 40-60% of the full length sequence of each of the above genes was present in the plasmids isolated from the genomic library. Therefore, primers were designed to amplify the full length of the respective genes and test their effects. The PCR products were inserted into empty pTrc99A plasmid and transformed into the XL1-Blue strain with the GFP-SsrA construct. Unfortunately, upon induction and analysis by flow cytometry, none of the full length genes resulted in higher fluorescence relative to the GFP-SsrA control, indicating that repression or inhibition of the Clp system did not occur. Additional isolated clones were checked by plate fluorescence and flow cytometry, and the genomic library was sorted a second time to isolate additional mutants. However, no additional clones showed an increase in fluorescence after the induction period.

The result indicates there are no gene fragments that upon multicopy expression serve as negative effectors of the Clp system. The increased fluorescence of clones obtained from library screening was apparently due to the overexpression of chimeric or incomplete proteins. Incomplete proteins are encoded in a genomic library since many clones contain pieces of genes which are not necessarily in the correct open-reading-frame. These polypeptides do not fold correctly and are thus likely to be targeted for degradation by the Clp system [39]. Since such proteins are overexpressed and would therefore be abundant, they compete with the GFP-SsrA fusion for degradation, increasing the fluorescence of the cell. Additionally, out-of-frame gene fragments may cause ribosome stalling during translation and recruitment of the SsrA RNA which tags the protein with the SsrA tag for degradation [32]. This process would also lead to polypeptides that would compete with the GFP-SsrA fusion for ClpXP. Such competition should lead to an increase in the fluorescence of the cell. In either case, the mechanism of ClpXP inhibition is trivial and does not merit further examination.

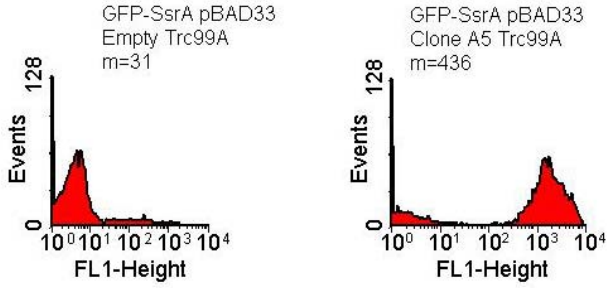

Figure 4.3. Example of a fluorescence histogram for a clone isolated from the genomic library. Clones isolated after back-transformation were highly fluorescent as compared to cells co-transformed with plasmid pGFP-SsrA and empty pTrc99A plasmid as a control. m, geometric mean of the fluorescence.

4.3.3 Selection for intragenic mutations in GFP-SsrA leading to higher cell fluorescence

4.3.3.A. GFP-SsrA library

Mutant versions of GFP that are more resistant to ClpXP degradation may be more resistant to mechanical unfolding since ClpX acts by pulling the protein from one end and imposing a mechanical force for unfolding the protein and feeding it into the proteolytic chamber. As the reporter protein, unfolded GFP is not fluorescent [31], thus any mutants with increased fluorescence over wild-type (WT) GFP would indicate clones that may be more resistant to mechanical unfolding.

The *gfp* gene was subjected to random mutagenesis using error-prone PCR [25]. The mutated *gfp* was then ligated into pGFP-SsrA replacing the WT *gfp* sequence. The ligation mixture was then electroporated into XL1-Blue to create a library with $2x10^6$ transformants. Five clones from the library were sequenced and averaged six mutations per *gfp* gene, giving an error rate of 0.7%.

After growth and induction of protein synthesis, the library was then screened by FACS to select highly fluorescent clones. $1.3x10^7$ events were sorted, and the top 0.03% fluorescent events, representing 3,455 cells, were collected. After growth overnight, 1,000 colonies were transferred to 96-well plates to be grown, replica plated, induced, and scanned by the fluorescence plate reader. The top ten fluorescent clones were grown overnight for plasmid purification and back-transformed into XL1-Blue. Flow cytometry showed that these clones were highly fluorescent five hours after induction (see Figure 4.4). However, upon sequencing the mutated *gfp* genes, it was discovered that all of the sequenced clones had a stop codon or nucleotide deletions before the SsrA tag. Thus higher fluorescence occurred due to the accumulation of cytoplasmic GFP since the SsrA tag was missing, and the protein was not targeted to degradation by ClpXP.

The fluorescence level of these clones is equivalent to that of cells expressing GFP without an SsrA tag. We hypothesized that any mutated GFP-SsrA would have a fluorescence much lower than that of cells expressing GFP alone. Therefore, an additional five clones with low fluorescence as analyzed by flow cytometry were selected and sequenced. Again, all of these five clones had stop codon changes before the SsrA tag.

The last two codons of *gfp* before the SsrA tag are 5'-TACAAA-3', and stop codons are encoded by the following codons: TAA, TAG and TGA. C→A or C→G substitutions in position 3 and an A→T in position 4 can yield a stop codon. Therefore, a new reverse primer that encodes these two codons was made so that no mutations could occur in them during error-prone PCR. A new library was constructed with this primer to yield $5.1x10^6$ transformants with an error rate of 0.3% or 2.1 nucleotide substitutions per gene based on the sequencing of nine clones. From this library, 4,311 cells were collected from a total of $2.2x10^7$ analyzed events representing the top 0.02% of the fluorescent events. 1,000 clones were grown and analyzed individually on the fluorescence plate reader, and the top ten were further analyzed by flow cytometry. Sequencing results indicated that two of these clones had stop codons upstream of the new primer and SsrA tag, and the other eight clones had 1 or 2 base pair deletions in the last 30 nucleotides of *gfp*, rendering the SsrA tag out-of-frame. Again, since the clones with the highest fluorescence had been examined and their fluorescence was similar to that of untagged GFP, the library was resorted, and clones with low fluorescence as compared to cells expressing intact GFP without an SsrA tag were selected. Analysis of $1.7x10^7$ events yielded 2,714 cells or 0.03% of the analyzed events. 500 clones were grown and analyzed individually on the fluorescence plate reader. The 26 most

fluorescent clones were analyzed by flow cytometry and sequenced. However, results again showed 10 stop codon errors and 16 base pair deletions in 26 sequenced clones.

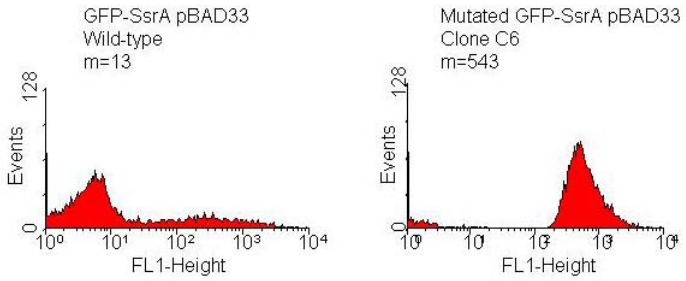

Figure 4.4. Example of a fluorescence histogram of a clone isolated from the mutated gfp library. Clones isolated after back-transformation were highly fluorescent as compared to WT GFP-SsrA. m, geometric mean of the fluorescence.

4.3.3.B. Placing the SsrA degradation tag at the N-terminus of GFP

In order to resolve this problem, I constructed a SsrA-GFP construct since previous work has shown that the Clp system can recognize and degrade proteins with N-terminal, C-terminal, or internal tags [40]. Since the SsrA tag is now upstream of GFP, any mutations in *gfp* that caused stop codons or nucleotide deletions to occur would not affect the SsrA tag. *E. coli* XL1-Blue cells were transformed with the *ssra-gfp* gene after cloning and verification by sequencing. Triplicate cultures were grown and induced at 0.1% and 0.2% arabinose induction levels, and samples were analyzed hourly by flow cytometry. As seen in Figure 4.5, the fluorescence did not decrease to baseline levels as seen with the GFP-SsrA fusion even for the lower induction level.

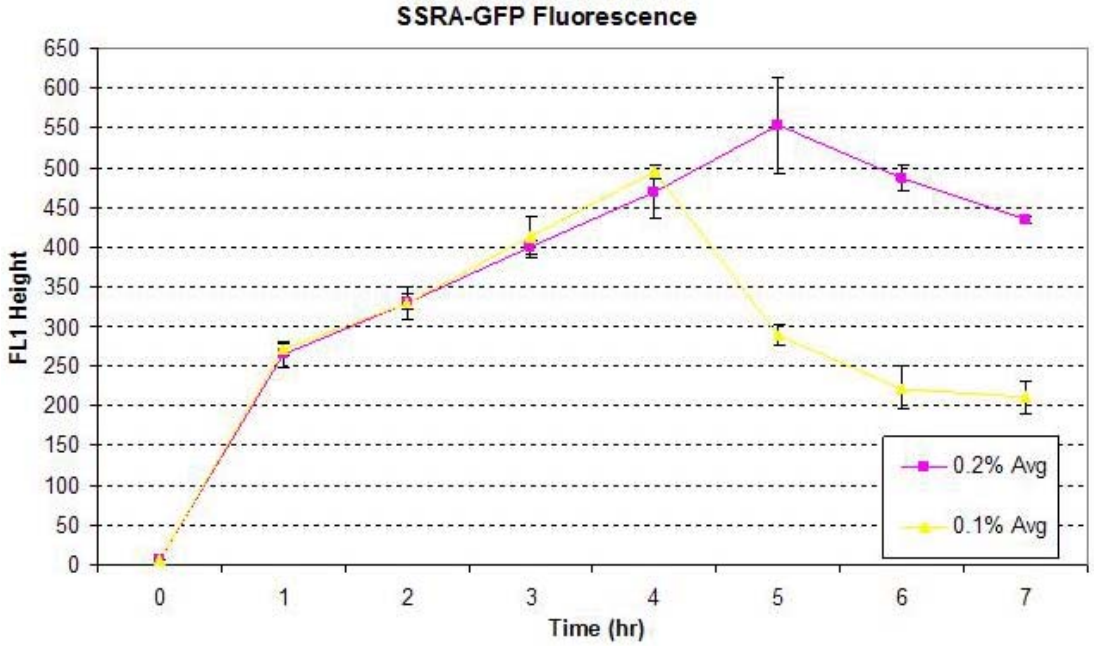

Figure 4.5. Time dependence of cell fluorescence in cells expressing SsrA-GFP. Flow cytometric analysis of SsrA-GFP in *E. coli* XL1-Blue for two different induction levels of arabinose, 0.2% and 0.1%. Error bars represent the standard deviation of the three samples per time point, and the line represents the average.

An attempt was made to alter the growth conditions of cells containing the SsrA-GFP fusion to reduce the high background fluorescence. The expression of SsrA-GFP in plasmid pSsrA-GFP is under control of the pBAD promoter. Protein synthesis is induced by arabinose, but it can also be repressed with glucose [24]. Cells with the pSsrA-GFP construct were grown using the conditions described above and induced to a 0.1% final arabinose concentration. However, after one hour of induction, the cultures were individually subjected to the following conditions: (A) cells from one culture were pelleted and resuspended and grown in fresh LB media with a concentration of 0.2% glucose, (B) a second culture had glucose added to achieve a final concentration of 0.2% glucose and (C) cells from a final aliquot were pelleted and resuspended and grown in fresh LB media. Figure 4.6 shows the results obtained for the three cultures. As

compared to the 0.1% arabinose culture in Figure 4.5, these cultures had lower maximum cell fluorescence and a lower level of baseline fluorescence except for the case of the culture with just the added glucose. The culture that was resuspended and regrown in fresh LB media had the mean fluorescence level of approximately 50.

Isolating positive clones over this background mean fluorescence would be difficult. As can be seen in Figure 4.7, there is considerable overlap between the culture fluorescence after 1 hour, when expression is at a maximum and after six hours. Because of the variance in each culture, it is not possible to distinguish between clones that have higher fluorescence due to physiological reasons and those that might express a mutated, more stable SsrA-GFP. In other words, the signal to noise is such that an unacceptably high number of false positive clones would be expected to arise during screening. Hence, a library was not constructed as isolating clones of interest from the false positives would be difficult and inefficient.

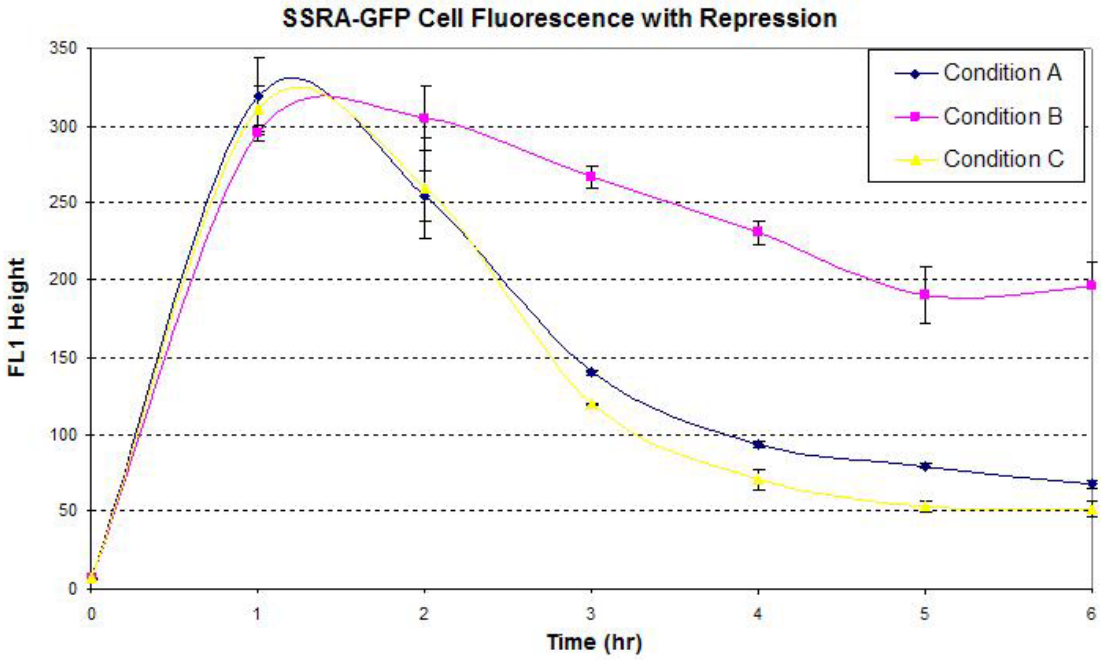

Figure 4.6. Time dependence of cell fluorescence in cells expressing SsrA-GFP with repression. Flow cytometric analysis of SsrA-GFP in *E. coli* XL1-Blue grown at 37°C for one hour before induction by 0.1% arabinose at t=0. At one hour following induction, the following conditions were implemented: (A) cells were pelleted and recultured in fresh LB media with a concentration of 0.2% glucose, (B) glucose was added to the culture to a concentration of 0.2% and (C) cells were pelleted and recultured into fresh LB media. Error bars represent the standard deviation of the three samples per time point, and the smoothed line curve represents the average.

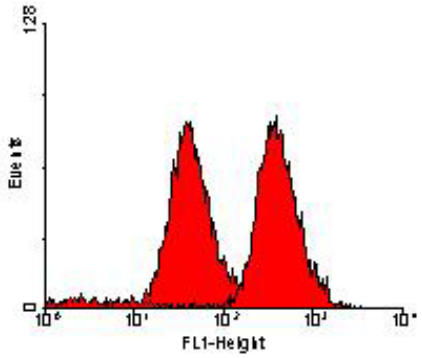

Figure 4.7. Overlay of histograms for cell culture in condition C**.** Histograms from flow cytometric analysis of cells that were pelleted and recultured into fresh LB media at one hour after induction (mean approximately 325) and six hours after induction (mean approximately 50).

4.3.3.C. GFP_{xho}-SsrA library

As discussed above, selection for the GFP-SsrA mutant conferring higher fluorescence yielded only stop codons and base pair deletions. To overcome these problems, a restriction site was engineered into the *gfp* gene. All of the residues in GFP are necessary for fluorescence except the last eight as determined from truncation studies [30, 31]. Therefore, only the N-terminus of the protein was subjected to mutagenesis. A restriction site was placed at a codon 52 amino acids from the C-terminus of GFP. Any mutation that introduces a stop codon within the first 186 residues of GFP should render the GFP non-fluorescent and reject it from being collected during sorting. Base pair deletions could allow the translation of almost full length protein, however, the codons after the deletion would be out-of-frame and as a result, the remainder of the polypeptide would encode an unrelated amino acid chain. Deletions in codons near the end of the polypeptide chain may still allow GFP to be fluorescent. Since the restriction site was placed at a codon distant from the C-terminus, this ensures that enough mutations following a base pair deletion would render GFP non-fluorescent. Flow cytometric analysis of the GFP_{xho}-SsrA fusion under the same growth and induction conditions as the GFP-SsrA fusion indicated that the introduced restriction site did not interfere with fluorescence of GFP as the time dependent fluorescence was equivalent to that shown in Figure 4.2.

Error-prone PCR was conducted on the first 558 nucleotides of *gfp*, and these fragments replaced the WT sequence of *gfp* upon ligation into $pGFP_{xho}$-SsrA. Transformation into *E. coli* XL1-Blue resulted in a library with $2x10^7$ transformants. Sequencing of ten clones indicated, on average, the presence of eight mutations per the first 558 nucleotides of *gfp* for an error rate of 1.4%. Following FACS screening of the library, 6,329 events representing the top 0.02% of fluorescence were isolated from a

total of $2.8x10^7$ events. It should be noted that the fluorescence of the library was equivalent to that of WT GFP_{xho}-SsrA, and no highly fluorescent events were observed. In order to collect the highest fluorescent events available, the lower end of the sort gate was set at a fluorescence of 20. Since the geometric mean of the library population was approximately 8, this small separation between the mean and the sort gate would result in the collection of WT clones, i.e. GFP_{xho}-SsrA with no mutations. Indeed, 500 colonies were scanned by fluorescence plate reader and all clones were equal or less fluorescent than the GFP_{xho}-SsrA parental protein.

Since analysis of this library during sorting indicated that no events had higher fluorescence than cells expressing GFP_{xho}-SsrA, it was possible that the high error rate used in the construction of the library introduced too many substitutions into the *gfp* gene, limiting the fraction of proteins that retained activity. Only the clones with few mutations would likely be fluorescent and because these clones would be statistically under-represented in a high error rate library, the probability of finding a clone with increased stability of GFP would be small. Hence, a new library with an error rate of 0.3% or 1.7 mutations per *gfp* gene was constructed. This library of $5x10^6$ transformants was sorted, and the top 0.001% fluorescent events, representing 263 out of $2.6x10^7$ total events, were collected. During sorting of this library, events with slightly higher fluorescence were detected so the lower end of the sort gate was set at a fluorescence of 40. However, all the collected events were equal or less fluorescent than the parental GFP_{xho}-SsrA fusion when analyzed by the fluorescence plate reader. In addition, several clones were randomly checked by flow cytometry, and all showed fluorescence equivalent to that of GFP_{xho}-SsrA (see Figure 4.8). In order to ensure that the expression of the mutant construct was equivalent to the WT construct, five clones were grown and protein synthesis was induced with arabinose. Cells from each clone culture along with a

WT culture were harvested after one hour of induction and the level of GFP protein was determined by Western blotting (see Figure 4.9). The Western blot indicated that expression of mutant GFP_{xho}-SsrA fusion in the clones was the same as for the cells expressing the parental protein. Therefore, no mutations that increased the stability of GFP against degradation by ClpXP, as indicated by the low fluorescence of the library, could be found.

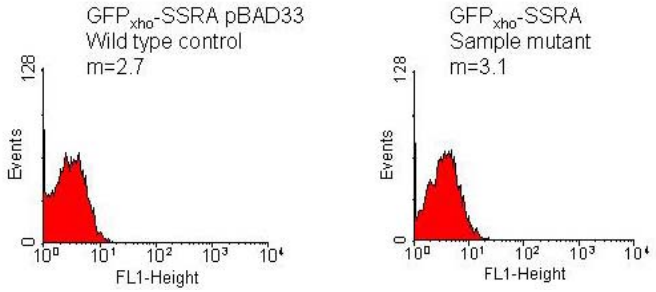

Figure 4.8. Fluorescence histograms from flow cytometric analysis of cells expressing a mutant GFPxho-SsrA collected from sorting and the WT control GFPxho-SsrA fusion. m, geometric mean of the fluorescence.

Figure 4.9. Western blot of GFPxho-SsrA and mutant proteins. Cells were grown for one hour at 37^{o}C in LB media before protein synthesis was induced with 0.2% arabinose. Cells were collected after one hour of induction from the WT control and five sample mutants. Lanes were loaded with an equivalent amount of cells as determined by optical cell density measurements.

4.3.3.D. GFP-RepA

The above studies indicate that random mutants of GFP-SsrA with increased stability against degradation by ClpXP could not be isolated. We therefore decided to investigate the possibility of using the same experimental process to isolate mutants with increased stability against ClpAP. ClpAP unfolds proteins in a manner similar to that of ClpXP. The RepA peptide tag targets proteins for degradation by ClpAP [41]. RepA is a molecular chaperone that initiates DNA replication from a plasmid that contains a P1 origin [42]. The first 15 amino acids of RepA contains the ClpA recognition signal which is sufficient to target a tagged protein to degradation by ClpAP [41].

A GFP-RepA fusion was constructed and expressed from plasmid pGFP-RepA. Cells were grown and induced under the same conditions used for the expression of GFP-SsrA, and samples were analyzed hourly by flow cytometry. As shown in Figure 4.10, the fluorescence of the GFP-RepA fusion barely increases over background, reaching a fluorescence height of 10 versus 200 for GFP-SsrA. Different levels of inducer were tested, but the maximum fluorescence never increased to a level that would allow clones to be isolated from a sort. This low maximum fluorescence could be due to low expression levels of the construct or perhaps higher degradation efficiency of GFP-RepA by ClpAP compared to the degradation of GFP-SsrA by ClpXP. No further experiments were conducted to examine expression levels since increasing expression levels high enough to achieve greater fluorescence would be difficult and could be toxic to the cell, in turn, creating growth and induction problems. Therefore, no library was fabricated from this construct.

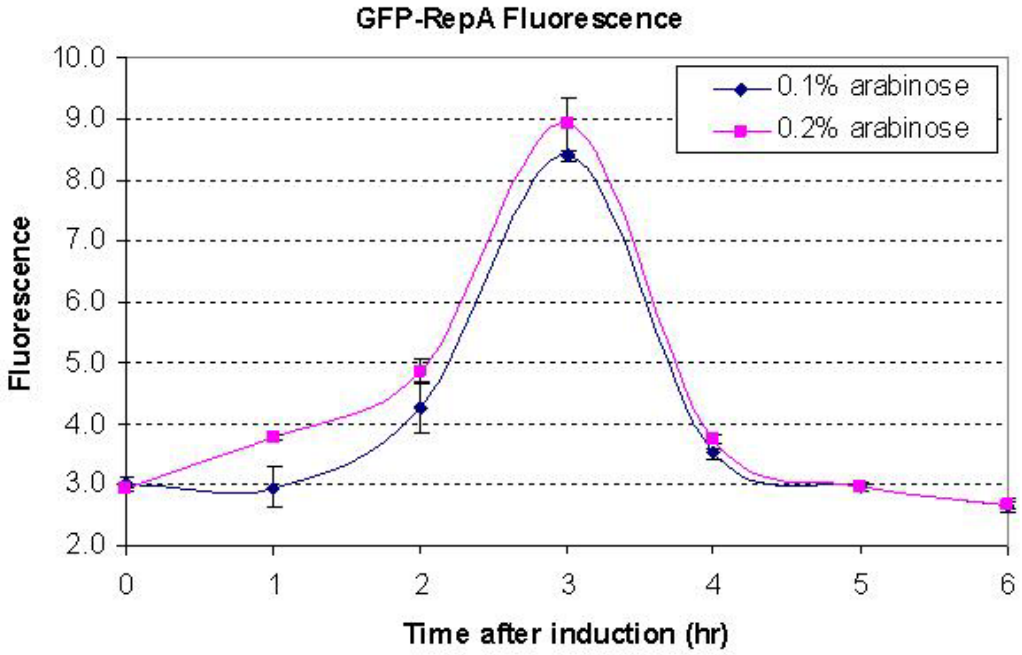

Figure 4.10. Time dependent cell fluorescence of GFP-RepA. Cultures were grown in triplicate for each induction level with conditions described previously. Samples were taken every hour and analyzed by flow cytometry. Error bars represent the standard deviation of the three samples per time point, and the smoothed line represents the average.

4.4 DISCUSSION

By using the GFP-SsrA fluorescent reporter protein fusion, we sought to discover multicopy genes or intragenic mutations that inhibit degradation by the ClpP proteolytic machinery of *E. coli*. Turning off or inhibiting the Clp system via the expression of a genomic DNA fragment would allow greater yields of protein to be produced since cytoplasmic proteolysis would be reduced. However, the Clp system must also retain its function to carry out its cellular roles during growth phases. Thus, inducible regulation of the activity of the Clp machinery could be important for recombinant protein production.

Unfortunately, screening of the genomic library by FACS did not yield any genes that showed reduced degradation of GFP-SsrA in the cell. Although the screen yielded several clones that conferred higher fluorescence, further examination revealed that they encoded only portions of genes that presumably gave rise to incomplete polypeptides that serves as competitive substrates for ClpXP degradation. The role of SsrA tagging is to direct incomplete polypeptides for degradation. Therefore, the finding that overexpression of incomplete polypeptides, such as those isolated from the genetic screen, inhibited degradation of GFP-SsrA by presumably serving as competitive inhibitors is not unexpected and is of little significance.

It should be noted that with the exception of SspB, other genetic studies have also failed to reveal regulators or inhibitors of the Clp system. The SspB protein acts as an adaptor protein for the ClpX chaperone [43]. The SspB protein recognizes the SsrA tag and tethers tagged proteins to ClpX, enhancing the rate of degradation [44]. Increasing the amount of SspB will increase the degradation rate of SsrA tagged proteins, but since ClpXP can degrade these tagged proteins even without SspB, it is not useful as a method of regulation [45].

We also searched for intragenic mutations in GFP that could decrease its degradation by ClpXP or ClpAP. Based on the mechanism of action of the ClpA and ClpX chaperones, mutations conferring resistance to degradation would be likely to exhibit greater stability to unfolding from with the N- or C-terminal end. By using the SsrA tag to target GFP to the ClpXP machinery, mutations could be introduced into *gfp* that would make it more resistant to degradation allowing differences in fluorescence to be used to isolate these mutants.

Initial experiments to isolate clones expressing GFP-SsrA and exhibiting higher fluorescence yielded false positive mutants in which the SsrA tag was missing or out-of-frame due to the introduction of stop codons or base pair deletions. The *Taq* polymerase used in the error-prone PCR to mutate the GFP gene has a normal error rate of approximately 1 error per 10,000 bases [46]. However, the basis of the error-prone PCR procedure is to increase the mutation rate, and these error rates are reported above for each library. Since the selection pressure of high fluorescence was applied and a large number of events were analyzed, these false positives would dominate the collected events, making the isolation of real positives difficult.

Changing the placement of the SsrA tag to the N-terminus or utilizing a different tag did not provide a useful strategy for screening due to the low fluorescence and the low signal-to-noise ratio that resulted. In the case of SsrA-GFP, the baseline fluorescence did not decrease to a low enough level to allow the discrimination of true positives from cells expressing unmutated SsrA-GFP. For GFP-RepA, the maximum fluorescence barely increased over baseline levels. In either case, the sort region would have to overlap the fluorescence distribution exhibited by cells expressing the respective unmutated protein thus leading to an unacceptably high frequency of false positives.

In a separate effort, a restriction site was engineered into *gfp* at a position corresponding to the codon 52 amino acids from its C-terminus. The introduction of this restriction site allowed the mutagenesis of only the first 558 of 714 nucleotides of *gfp* in the GFP-SsrA fusion. This strategy eliminated false positives due to stop codons or deletions since such mutants would not produce fluorescent GFP. Screening of two libraries with different error rates, however, did not yield any clones with increased fluorescence over WT GFP. This indicated that no mutations could stabilize GFP against degradation by ClpXP while still maintaining its fluorescence.

Since the completion of these studies, new insights have been discovered about the mechanism of protein stability to degradation by the Clp machinery. By using different mutants of the same protein, the rate of degradation by ClpXP of these proteins was shown not to depend on the global stability of the protein, as determined by the free energy of unfolding [47]. Rather, the local structure of the protein adjacent to the tag influenced the rate of degradation and ATP consumption during the degradation process [15]. Furthermore, it was determined that proteins with an alpha helix or irregular loops next to the degradation tag were unfolded more easily than proteins that had beta strands next to the tag [15, 48].

In light of these findings, our approach to change the overall stability of GFP would then have had no effect on increasing its resistance to degradation against ClpXP. In addition, since GFP is an all beta-sheet protein, its resistance to degradation may already be at a maximum. Bohn *et al.* found that GFP-SsrA could not be stabilized against ClpXP as well with the use of mutational insertions [49]. Also, resistance to degradation could not be achieved by screening a genomic library to find chaperones that would stabilize GFP-SsrA against ClpXP [49]. Finally, using transposon mutagenesis,

Bohn *et al.* found that only inactivation of *clpX* or *clpP* resulted in stabilization of GFP-SsrA [49]. This independent research confirms the results of this study.

4.5 REFERENCES

1. Jenal, U. and R. Hengge-Aronis, *Regulation by proteolysis in bacterial cells.* Curr Opin Microbiol, 2003. **6**(2): p. 163-72.
2. Straus, D.B., W.A. Walter, and C.A. Gross, *Escherichia coli heat shock gene mutants are defective in proteolysis.* Genes Dev, 1988. **2**(12B): p. 1851-8.
3. Porankiewicz, J., J. Wang, and A.K. Clarke, *New insights into the ATP-dependent Clp protease: Escherichia coli and beyond.* Mol Microbiol, 1999. **32**(3): p. 449-58.
4. Jenal, U. and T. Fuchs, *An essential protease involved in bacterial cell-cycle control.* Embo J, 1998. **17**(19): p. 5658-69.
5. Lee, C., et al., *ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal.* Mol Cell, 2001. **7**(3): p. 627-37.
6. Zellmeier, S., W. Schumann, and T. Wiegert, *Involvement of Clp protease activity in modulating the Bacillus subtilissigmaw stress response.* Mol Microbiol, 2006. **61**(6): p. 1569-82.
7. Thibault, G., et al., *Large nucleotide-dependent movement of the N-terminal domain of the ClpX chaperone.* Embo J, 2006. **25**(14): p. 3367-76.
8. Szyk, A. and M.R. Maurizi, *Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site.* J Struct Biol, 2006.
9. Singh, S.K., et al., *Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP.* Proc Natl Acad Sci U S A, 2000. **97**(16): p. 8898-903.
10. Thompson, M.W. and M.R. Maurizi, *Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates.* J Biol Chem, 1994. **269**(27): p. 18201-8.
11. Gribun, A., et al., *The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation.* J Biol Chem, 2005. **280**(16): p. 16185-96.
12. Banecki, B., et al., *Structure-function analysis of the zinc-binding region of the Clpx molecular chaperone.* J Biol Chem, 2001. **276**(22): p. 18843-8.
13. Smith, C.K., T.A. Baker, and R.T. Sauer, *Lon and Clp family proteases and chaperones share homologous substrate-recognition domains.* Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6678-82.
14. Burton, R.E., et al., *Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine.* Embo J, 2001. **20**(12): p. 3092-100.
15. Kenniston, J.A., et al., *Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation.* J Struct Biol, 2004. **146**(1-2): p. 130-40.
16. Paul, R., et al., *Shear stress related blood damage in laminar couette flow.* Artif Organs, 2003. **27**(6): p. 517-29.

17. Long, M., S. Lu, and G. Sun, *Kinetics of receptor-ligand interactions in immune responses.* Cell Mol Immunol, 2006. **3**(2): p. 79-86.
18. Chan, B.P., W.M. Reichert, and G.A. Truskey, *Effect of streptavidin RGD mutant on the adhesion of endothelial cells.* Biotechnol Prog, 2004. **20**(2): p. 566-75.
19. Chada, D., T. Mather, and M.U. Nollert, *The Synergy Site of Fibronectin Is Required for Strong Interaction with the Platelet Integrin alphaIIbbeta3.* Ann Biomed Eng, 2006.
20. Cumbers, S.J., et al., *Generation and iterative affinity maturation of antibodies in vitro using hypermutating B-cell lines.* Nat Biotechnol, 2002. **20**(11): p. 1129-34.
21. Baca, M., et al., *Antibody humanization using monovalent phage display.* J Biol Chem, 1997. **272**(16): p. 10678-84.
22. Chandu, D. and D. Nandi, *Comparative genomics and functional roles of the ATP-dependent proteases Lon and Clp during cytosolic protein degradation.* Res Microbiol, 2004. **155**(9): p. 710-9.
23. Voges, D., P. Zwickl, and W. Baumeister, *The 26S proteasome: a molecular machine designed for controlled proteolysis.* Annu Rev Biochem, 1999. **68**: p. 1015-68.
24. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.
25. Fromant, M., S. Blanquet, and P. Plateau, *Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction.* Anal Biochem, 1995. **224**(1): p. 347-53.
26. Levy, R., et al., *Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones.* Protein Expr Purif, 2001. **23**(2): p. 338-47.
27. Thomas, J.D., et al., *Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli.* Mol Microbiol, 2001. **39**(1): p. 47-53.
28. Chae, H.J., et al., *Framework for online optimization of recombinant protein expression in high-cell-density Escherichia coli cultures using GFP-fusion monitoring.* Biotechnol Bioeng, 2000. **69**(3): p. 275-85.
29. Feilmeier, B.J., et al., *Green fluorescent protein functions as a reporter for protein localization in Escherichia coli.* J Bacteriol, 2000. **182**(14): p. 4068-76.
30. Tsien, R.Y., *The green fluorescent protein.* Annu Rev Biochem, 1998. **67**: p. 509-44.
31. Zimmer, M., *Green fluorescent protein (GFP): applications, structure, and related photophysical behavior.* Chem Rev, 2002. **102**(3): p. 759-81.
32. Karzai, A.W., E.D. Roche, and R.T. Sauer, *The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue.* Nat Struct Biol, 2000. **7**(6): p. 449-55.
33. Keiler, K.C., P.R. Waller, and R.T. Sauer, *Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA.* Science, 1996. **271**(5251): p. 990-3.
34. DeLisa, M.P., et al., *Genetic analysis of the twin arginine translocator secretion pathway in bacteria.* J Biol Chem, 2002. **277**(33): p. 29825-31.

35. Brown, T.A., *Genomes*. 2002, New York: Garland Science.
36. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.
37. Blattner, F.R., et al., *The complete genome sequence of Escherichia coli K-12.* Science, 1997. **277**(5331): p. 1453-74.
38. Leipold, R.J. and P. Dhurjati, *Specialized ribosomes in Escherichia coli.* Biotechnol Prog, 1993. **9**(5): p. 443-9.
39. Thomsen, L.E., et al., *ClpP is involved in the stress response and degradation of misfolded proteins in Salmonella enterica serovar Typhimurium.* Microbiology, 2002. **148**(Pt 9): p. 2727-33.
40. Hoskins, J.R., et al., *ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11037-42.
41. Hoskins, J.R. and S. Wickner, *Two peptide sequences can function cooperatively to facilitate binding and unfolding by ClpA and degradation by ClpAP.* Proc Natl Acad Sci U S A, 2006. **103**(4): p. 909-14.
42. Wickner, S., et al., *A molecular chaperone, ClpA, functions like DnaK and DnaJ.* Proc Natl Acad Sci U S A, 1994. **91**(25): p. 12218-22.
43. Wah, D.A., et al., *Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer.* Chem Biol, 2002. **9**(11): p. 1237-45.
44. Song, H.K. and M.J. Eck, *Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine.* Mol Cell, 2003. **12**(1): p. 75-86.
45. McGinness, K.E., T.A. Baker, and R.T. Sauer, *Engineering controllable protein degradation.* Mol Cell, 2006. **22**(5): p. 701-7.
46. Cariello, N.F., J.A. Swenberg, and T.R. Skopek, *Fidelity of Thermococcus litoralis DNA polymerase (Vent) in PCR determined by denaturing gradient gel electrophoresis.* Nucleic Acids Res, 1991. **19**(15): p. 4193-8.
47. Kenniston, J.A., et al., *Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine.* Cell, 2003. **114**(4): p. 511-20.
48. Matouschek, A., *Protein unfolding--an important process in vivo?* Curr Opin Struct Biol, 2003. **13**(1): p. 98-109.
49. Bohn, C., E. Binet, and P. Bouloc, *Screening for stabilization of proteins with a trans-translation signature in Escherichia coli selects for inactivation of the ClpXP protease.* Mol Genet Genomics, 2002. **266**(5): p. 827-31.

Chapter 5

Conclusions and Recommendations

In this work, we applied the high throughput method of fluorescence activated cell sorting (FACS) to investigate protein export and engineering. Using directed evolution to mutate the genes of interest and then linking a cell's genotype to a fluorescent phenotype, FACS allowed us to screen millions of mutants quickly and efficiently.

We have used FACS to investigate problems related to protein export via the Tat pathway and protein stability to mechanical unfolding and resistance to proteolysis by the Clp system of *E. coli*.

We examined ways to increase recombinant protein yields in the *E. coli* periplasm through the Tat pathway. The Tat pathway also offers a quality control advantage in that it exports only correctly folded proteins. This is especially useful for antibody fragment production since the Tat pathway will ensure that only correctly folded fragments in which disulfide bond formation has occurred are transported to the periplasm in an oxidizing cytoplasm strain. Therefore, only the periplasmic fraction of the cells needs to be harvested and only fully functional antibody fragments would be obtained.

Since the efficiency of the Tat pathway is low, we used FACS to screen mutants of the anti-digoxin 26-10 scFv that exhibit improved Tat export. We isolated a mutant scFv C8 that is produced at an increased periplasmic yield of approximately two-fold over the WT scFv. Further analysis of the clone revealed that the mutations increased the rate of refolding 2.5 times that of the WT 26-10. This faster refolding rate likely keeps

the mutant protein from participating in off-pathway reactions such as protein aggregation, thereby allowing more competent protein to be available for Tat export.

However, since this improved yield is still lower than that achieved when export is targeted to the Sec pathway, additional experiments were conducted to further increase the yield. Using various *E. coli* mutants having different mutations that render the cytoplasm oxidizing, as required for disulfide bond formation, or changing the growth conditions did not increase the fluorescence. By co-expressing different proteins that could alleviate the saturation of the Tat pathway or provide chaperone functions, we determined that co-expression of the Tat translocon genes yield the largest improvement in cell fluorescence and therefore export yield. Overall, co-expression of the TatABC proteins with the C8 mutant scFv resulted in an increase in fluorescence that was five times greater than the fluorescence obtained by exporting the parental anti-digoxin scFv. Even so, this maximum fluorescence was only 16-fold over background which is still approximately half that achieved through the Sec pathway. Thus, additional methods of improvement may need to be investigated in order to increase the level of Tat export such as engineering a better strain based on the FA113 strain.

In other studies, we sought to discover any regulators or inhibitors of the Clp system present in *E. coli* utilizing the GFP-SsrA fluorescent reporter protein fusion. Finding novel regulators of the Clp system, apart from its biological significance could help suppress proteolysis and thus allow increased yields of protein to be produced in *E. coli*. However, screening of the genomic library did not yield any genes that played a role in the regulation or inhibition of the Clp system. This is most likely due to the fact that the Clp system is used to regulate many cellular processes, and its down-regulation may be detrimental.

The same fluorescent construct was used to explore the engineering of more resistant proteins to mechanical forces and possibly to shear. Since the Clp system unfolds proteins by applying a mechanical force, determining ways to increase a protein's stability against Clp degradation may correlate with a higher resistance to shear. However, after several different libraries were constructed and screened, no mutations were found that could stabilize GFP against the Clp system. Research conducted after this study found that resistance to degradation by the Clp system depends on local stability near the degradation tag rather than the global stability of the protein. Hence, although mutations in *gfp* could perhaps increase the global stability of the protein, they would be unlikely to increase the protein's resistance to shear, given that GFP folds into a highly cooperative beta-barrel.

5.1 Recommendations for future studies

5.1.1 Reducing ClpXP degradation

Although the screening of a genomic library did not yield any genes which inhibit the ClpXP machinery, other approaches could be tried to reduce proteolysis. As discussed in Chapter 1, the SspB protein has a C-terminal XB peptide motif that is involved in flexible tethering to ClpX. The SspB adaptor recognizes the tag then dimerizes with another SspB adaptor to deliver the tagged substrate to ClpX using the two XB peptide motifs to anchor the dimerized adaptor to ClpX.

By altering the residues of SspB that recognize the SsrA tag, a mutant SspB protein could be engineered which only binds but does not translocate the tag. This would allow the SspB adaptor to recognize a substrate, dimerize and associate with ClpXP. By not being able to translocate the tag, the SspB dimmer would effectively be blocking that ClpXP molecule from carrying out degradation. However, since the association and disassociation of the SspB dimer to ClpXP is an equilibrium process and ClpXP can degrade tagged substrates in the absence of SspB, it will also be necessary to evolve the XB peptide motifs and flexible linker that tethers the SspB dimer to ClpXP. By increasing the affinity of the SspB dimmer to ClpXP, overexpression of this mutant SspB could effectively block all the constitutively expressed ClpXP molecules. As this may affect cell viability and growth, having the mutant SspB on an inducible plasmid may allow it to be overexpressed only when the recombinant protein of interested is expressed.

This proposed study could be conducted in the same manner as the completed study in Chapter 4. Use of the GFP-SsrA construct co-expressed with mutant SspB libraries would allow the use of FACS as a high throughput screening method. This

project can be done in two stages by creation of a library of SspB mutants in which amino acid substitutions in the SsrA recognition site have been made followed by another library in which amino acid substitutions have been made in the flexible tether of the most fluorescent mutant SspB proteins isolated from the first library. On the other hand, one library could be constructed where mutations are made in both the SsrA tag recognition site and the flexible tether. Either way, mutants of SspB that only bind the SsrA tag and associate with ClpXP with a higher affinity would have higher fluorescence since less GFP-SsrA would be degraded over time than WT versions of SspB. This would allow rapid and efficient screening of the mutants by FACS and would yield a means to selectively inhibit the ClpXP system.

5.1.2 Further increasing Tat export efficiency

As seen in Chapter 3, co-expression of TatABC with the C8 ssTorA-26-10 scFv resulted in a five-fold increase over WT ssTorA-26-10 scFv expressed alone. However, in order to match the export efficiency of the Sec pathway, further improvements must be made. Only then can the Tat pathway become a viable means of recombinant protein production for biotechnology.

Since co-expression of all the various proteins examined in Chapter 3 produced an increase in fluorescence, it may be possible to place some of these proteins under inducible control on the *E. coli* chromosome. This would allow additional proteins to be co-expressed with the C8 ssTorA-26-10 scFv without having to contend with plasmid incompatibility or plasmid loss. For example, a FA113 strain in which the TorD and TatABC genes were placed under inducible control was created. However, upon induction of these genes and the WT ssTorA-26-10 scFv no fluorescence improvement was observed. This was most likely due to the incorrect placement of the genes on the

chromosome or the introduction of errors into these genes during the genetic recombination process used to create the new strain. Therefore, more testing and troubleshooting of the strain will be necessary.

When a new strain of FA113 is successfully created, the inducible combination of TorD and TatABC on the chromosome would allow the co-expression of another protein from a plasmid like PspA. Although individually TorD, TatABC and PspA increased the fluorescence of the C8 scFv 2.4, 1.3- and 1.5-fold, respectively, their combination most likely will not result in a total cumulative gain since the Tat pathway may reach a maximum export rate. However, their combination may improve the C8 scFv fluorescence 3- to 3.5-fold over the C8 scFv fluorescence alone. With the two-fold improvement of the C8 scFv over WT, this total improvement could increase protein expression seven times over WT leading to a purified yield of 0.91 mg/L/OD based on the yield values from Chapter 3. This value is significantly closer to the 1 g/L/OD values obtained for proteins exported via Sec.

Finally, the new strain with the C8 ssTorA-26-10 construct will need to be cultured in a stirred tank bioreactor to determine the actual yields that can obtained. This approach would allow for better mixing and oxygen transfer rates than the shake flasks used in the mid-scale expression experiments in Chapter 2. The growth conditions can also be optimized again since growth in these experiments in Chapter 3 may have been oxygen limited. Also, the implementation of pH control on the bioreactor would improve cell viability by maintaining a pH that is optimal for *E. coli* strain growth. These improvements may increase cell densities allowing an even greater yield of 26-10 scFv per volume and show that the Tat pathway can indeed be utilized for recombinant protein production.

Author's Publication List

Tullman-Ercek, D., DeLisa, M. P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T., Georgiou, G. (Submitted). "Export pathway selectivity of *Escherichia coli* twin-arginine translocation signal peptides."

Ribnicky, B., Van Blarcom, T., Georgiou, G. (Submitted). "A scFv Antibody Mutant Isolated in a Genetic Screen for Improved Export via the Twin Arginine Transporter Pathway Exhibits Faster Folding."

Bibliography

Ades, S. E. (2004). "Proteolysis: Adaptor, adaptor, catch me a catch." Curr Biol **14**(21): R924-6.

Ades, S. E. (2006). "AAA+ molecular machines: firing on all cylinders." Curr Biol **16**(2): R46-8.

Aharoni, A., A. D. Griffiths, et al. (2005). "High-throughput screens and selections of enzyme-encoding genes." Curr Opin Chem Biol **9**(2): 210-6.

Alami, M., I. Luke, et al. (2003). "Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli." Mol Cell **12**(4): 937-46.

Alberts, B., A. Johnson, et al. (2002). Molecular Biology of the Cell.

Alder, N. N. and S. M. Theg (2003). "Energy use by biological protein transport pathways." Trends Biochem Sci **28**(8): 442-51.

Allen, S. C., C. M. Barrett, et al. (2002). "Essential cytoplasmic domains in the Escherichia coli TatC protein." J Biol Chem **277**(12): 10362-6.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10.

Arbabi-Ghahroudi, M., J. Tanha, et al. (2005). "Prokaryotic expression of antibodies." Cancer Metastasis Rev **24**(4): 501-19.

Arnold, F. and G. Georgiou (2003). Directed enzyme evolution: Screening and selection methods. Totowa, NJ.

Baca, M., L. G. Presta, et al. (1997). "Antibody humanization using monovalent phage display." J Biol Chem **272**(16): 10678-84.

Bailey, A. M., M. A. Webber, et al. (2006). "Medium plays a role in determining expression of acrB, marA, and soxS in Escherichia coli." Antimicrob Agents Chemother **50**(3): 1071-4.

Banecki, B., A. Wawrzynow, et al. (2001). "Structure-function analysis of the zinc-binding region of the Clpx molecular chaperone." J Biol Chem **276**(22): 18843-8.

Baneyx, F. (1999). *"In vivo Folding of Recombinant proteins in Escherichia coli"* in Manual of Industrial Microbiology and Biotechnology, ASM Press.

Baneyx, F. and M. Mujacic (2004). "Recombinant protein folding and misfolding in Escherichia coli." Nat Biotechnol **22**(11): 1399-408.

Barrett, C. M., N. Ray, et al. (2003). "Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in Escherichia coli." Biochem Biophys Res Commun **304**(2): 279-84.

Barrett, C. M. and C. Robinson (2005). "Evidence for interactions between domains of TatA and TatB from mutagenesis of the TatABC subunits of the twin-arginine translocase." Febs J **272**(9): 2261-75.

Barth, S., M. Huhn, et al. (2000). "Compatible-solute-supported periplasmic expression of functional recombinant proteins under stress conditions." Appl Environ Microbiol **66**(4): 1572-9.

Berks, B. C., T. Palmer, et al. (2003). "The Tat protein translocation pathway and its role in microbial physiology." Adv Microb Physiol **47**: 187-254.

Berks, B. C., F. Sargent, et al. (2000). "The Tat protein export pathway." Mol Microbiol **35**(2): 260-74.

Bessette, P. H., F. Aslund, et al. (1999). "Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm." Proc Natl Acad Sci U S A **96**(24): 13703-8.
Bessonneau, P., V. Besson, et al. (2002). "The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure." Embo J **21**(5): 995-1003.
Biosciences, B. (2004). "BD FACSCalibur Flow Cytometer." from www.bdbiosciences.com.
Blaber, M., X. J. Zhang, et al. (1993). "Structural basis of amino acid alpha helix propensity." Science **260**(5114): 1637-40.
Blattner, F. R., G. Plunkett, 3rd, et al. (1997). "The complete genome sequence of Escherichia coli K-12." Science **277**(5331): 1453-74.
Blaudeck, N., P. Kreutzenbeck, et al. (2005). "Isolation and characterization of bifunctional Escherichia coli TatA mutant proteins that allow efficient tat-dependent protein translocation in the absence of TatB." J Biol Chem **280**(5): 3426-32.
Bochtler, M., C. Hartmann, et al. (2000). "The structures of HsIU and the ATP-dependent protease HsIU-HsIV." Nature **403**(6771): 800-5.
Boder, E. T., K. S. Midelfort, et al. (2000). "Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity." Proc Natl Acad Sci U S A **97**(20): 10701-5.
Bohn, C., E. Binet, et al. (2002). "Screening for stabilization of proteins with a trans-translation signature in Escherichia coli selects for inactivation of the ClpXP protease." Mol Genet Genomics **266**(5): 827-31.
Bolon, D. N., R. A. Grant, et al. (2004). "Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease." Mol Cell **16**(3): 343-50.
Breyton, C., W. Haase, et al. (2002). "Three-dimensional structure of the bacterial protein-translocation complex SecYEG." Nature **418**(6898): 662-5.
Brissette, J. L., M. Russel, et al. (1990). "Phage shock protein, a stress protein of Escherichia coli." Proc Natl Acad Sci U S A **87**(3): 862-6.
Brown, T. A. (2002). Genomes. New York, Garland Science.
Burton, B. M. and T. A. Baker (2005). "Remodeling protein complexes: insights from the AAA+ unfoldase ClpX and Mu transposase." Protein Sci **14**(8): 1945-54.
Burton, R. E., S. M. Siddiqui, et al. (2001). "Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine." Embo J **20**(12): 3092-100.
Cabantous, S., J. D. Pedelacq, et al. (2005). "Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from Mycobacterium tuberculosis." J Struct Funct Genomics **6**(2-3): 113-9.
Cadwell, R. C. and G. F. Joyce (1994). "Mutagenic PCR." PCR Methods Appl **3**(6): S136-40.
Cariello, N. F., J. A. Swenberg, et al. (1991). "Fidelity of Thermococcus litoralis DNA polymerase (Vent) in PCR determined by denaturing gradient gel electrophoresis." Nucleic Acids Res **19**(15): 4193-8.

Carmel-Harel, O. and G. Storz (2000). "Roles of the glutathione- and thioredoxin-dependent reduction systems in the Escherichia coli and saccharomyces cerevisiae responses to oxidative stress." Annu Rev Microbiol **54**: 439-61.
Carter, P. J. (2006). "Potent antibody therapeutics by design." Nat Rev Immunol **6**(5): 343-57.
Chada, D., T. Mather, et al. (2006). "The Synergy Site of Fibronectin Is Required for Strong Interaction with the Platelet Integrin alphaIIbbeta3." Ann Biomed Eng.
Chae, H. J., M. P. Delisa, et al. (2000). "Framework for online optimization of recombinant protein expression in high-cell-density Escherichia coli cultures using GFP-fusion monitoring." Biotechnol Bioeng **69**(3): 275-85.
Chan, B. P., W. M. Reichert, et al. (2004). "Effect of streptavidin RGD mutant on the adhesion of endothelial cells." Biotechnol Prog **20**(2): 566-75.
Chanal, A., C. L. Santini, et al. (2003). "Specific inhibition of the translocation of a subset of Escherichia coli TAT substrates by the TorA signal peptide." J Mol Biol **327**(3): 563-70.
Chandu, D. and D. Nandi (2004). "Comparative genomics and functional roles of the ATP-dependent proteases Lon and Clp during cytosolic protein degradation." Res Microbiol **155**(9): 710-9.
Chen, G., I. Dubrawsky, et al. (1999). "In vitro scanning saturation mutagenesis of all the specificity determining residues in an antibody binding site." Protein Eng **12**(4): 349-56.
Chen, G., A. Hayhurst, et al. (2001). "Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS)." Nat Biotechnol **19**(6): 537-42.
Chen, J., J. L. Song, et al. (1999). "Chaperone activity of DsbC." J Biol Chem **274**(28): 19601-5.
Choi, K. H. and S. Licht (2005). "Control of peptide product sizes by the energy-dependent protease ClpAP." Biochemistry **44**(42): 13921-31.
Clark, S. A. and S. M. Theg (1997). "A folded protein can be transported across the chloroplast envelope and thylakoid membranes." Mol Biol Cell **8**(5): 923-34.
Clarke, J. and A. R. Fersht (1993). "Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation." Biochemistry **32**(16): 4322-9.
Cline, K. and H. Mori (2001). "Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport." J Cell Biol **154**(4): 719-29.
Connolly, T. and R. Gilmore (1989). "The signal recognition particle receptor mediates the GTP-dependent displacement of SRP from the signal sequence of the nascent polypeptide." Cell **57**(4): 599-610.
Cristobal, S., J. W. de Gier, et al. (1999). "Competition between Sec- and TAT-dependent protein translocation in Escherichia coli." Embo J **18**(11): 2982-90.
Cumbers, S. J., G. T. Williams, et al. (2002). "Generation and iterative affinity maturation of antibodies in vitro using hypermutating B-cell lines." Nat Biotechnol **20**(11): 1129-34.

Cupit, P. M., J. A. Whyte, et al. (1999). "Cloning and expression of single chain antibody fragments in Escherichia coli and Pichia pastoris." Lett Appl Microbiol **29**(5): 273-7.
Daugherty, P. S., G. Chen, et al. (2000). "Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies." Proc Natl Acad Sci U S A **97**(5): 2029-34.
Daugherty, P. S., G. Chen, et al. (1998). "Antibody affinity maturation using bacterial surface display." Protein Eng **11**(9): 825-32.
De Keersmaeker, S., K. Vrancken, et al. (2006). "Evaluation of TatABC overproduction on Tat- and Sec-dependent protein secretion in Streptomyces lividans." Arch Microbiol.
de Leeuw, E., T. Granjon, et al. (2002). "Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes." J Mol Biol **322**(5): 1135-46.
De Leeuw, E., I. Porcelli, et al. (2001). "Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway." FEBS Lett **506**(2): 143-8.
De Lorenzo, F., R. F. Goldberger, et al. (1966). "Purification and properties of an enzyme from beef liver which catalyzes sulfhydryl-disulfide interchange in proteins." J Biol Chem **241**(7): 1562-7.
DeLisa, M. P., P. Lee, et al. (2004). "Phage shock protein PspA of Escherichia coli relieves saturation of protein export via the Tat pathway." J Bacteriol **186**(2): 366-73.
DeLisa, M. P., P. Samuelson, et al. (2002). "Genetic analysis of the twin arginine translocator secretion pathway in bacteria." J Biol Chem **277**(33): 29825-31.
DeLisa, M. P., D. Tullman, et al. (2003). "Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway." Proc Natl Acad Sci U S A **100**(10): 6115-20.
Doyle, S. M., E. Anderson, et al. (2003). "Rapid unfolding of a domain populates an aggregation-prone intermediate that can be recognized by GroEL." J Mol Biol **332**(4): 937-51.
Drews, J. (2000). "Drug discovery: a historical perspective." Science **287**(5460): 1960-4.
Drews, J. (2003). "Stategic trends in the drug industry." Drug Discov Today **8**(9): 411-20.
Eijsink, V. G., A. Bjork, et al. (2004). "Rational engineering of enzyme stability." J Biotechnol **113**(1-3): 105-20.
Eijsink, V. G., S. Gaseidnes, et al. (2005). "Directed evolution of enzyme stability." Biomol Eng **22**(1-3): 21-30.
Emmerson, P. J., S. Archer, et al. (1997). "Synthesis and characterization of 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (BODIPY)-labeled fluorescent ligands for the mu opioid receptor." Biochem Pharmacol **54**(12): 1315-22.
Farrell, C. M., A. D. Grossman, et al. (2005). "Cytoplasmic degradation of ssrA-tagged proteins." Mol Microbiol **57**(6): 1750-61.
Feilmeier, B. J., G. Iseminger, et al. (2000). "Green fluorescent protein functions as a reporter for protein localization in Escherichia coli." J Bacteriol **182**(14): 4068-76.

Feldhaus, M. J., R. W. Siegel, et al. (2003). "Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library." Nat Biotechnol **21**(2): 163-70.
Fernandes, A. P. and A. Holmgren (2004). "Glutaredoxins: glutathione-dependent redox enzymes with functions far beyond a simple thioredoxin backup system." Antioxid Redox Signal **6**(1): 63-74.
Fisher, A. C., W. Kim, et al. (2006). "Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway." Protein Sci **15**(3): 449-58.
Flynn, J. M., S. B. Neher, et al. (2003). "Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals." Mol Cell **11**(3): 671-83.
Fromant, M., S. Blanquet, et al. (1995). "Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction." Anal Biochem **224**(1): 347-53.
Gai, D., R. Zhao, et al. (2004). "Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen." Cell **119**(1): 47-60.
Genest, O., F. Seduk, et al. (2006). "Signal peptide protection by specific chaperone." Biochem Biophys Res Commun **339**(3): 991-5.
Georgiou, G. (2000). "Analysis of large libraries of protein mutants using flow cytometry." Adv Protein Chem **55**: 293-315.
Georgiou, G. and L. Segatori (2005). "Preparative expression of secreted proteins in bacteria: status report and future prospects." Curr Opin Biotechnol **16**(5): 538-45.
Glockshuber, R., M. Malia, et al. (1990). "A comparison of strategies to stabilize immunoglobulin Fv-fragments." Biochemistry **29**(6): 1362-7.
Gohlke, U., L. Pullan, et al. (2005). "The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter." Proc Natl Acad Sci U S A **102**(30): 10482-6.
Gottesman, S. (1996). "Proteases and their targets in Escherichia coli." Annu Rev Genet **30**: 465-506.
Gottesman, S. (2003). "Proteolysis in bacterial regulatory circuits." Annu Rev Cell Dev Biol **19**: 565-87.
Gottesman, S., W. P. Clark, et al. (1993). "ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities." J Biol Chem **268**(30): 22618-26.
Gottesman, S. and M. R. Maurizi (1992). "Regulation by proteolysis: energy-dependent proteases and their targets." Microbiol Rev **56**(4): 592-621.
Gottesman, S., E. Roche, et al. (1998). "The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system." Genes Dev **12**(9): 1338-47.
Gottesman, S., S. Wickner, et al. (1997). "Protein quality control: triage by chaperones and proteases." Genes Dev **11**(7): 815-23.
Gough, K. C., Y. Li, et al. (1999). "Selection of phage antibodies to surface epitopes of Phytophthora infestans." J Immunol Methods **228**(1-2): 97-108.
Gribun, A., M. S. Kimber, et al. (2005). "The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form

flexible loops that are essential for ClpXP and ClpAP complex formation." J Biol Chem **280**(16): 16185-96.
Grimaud, R., M. Kessel, et al. (1998). "Enzymatic and structural similarities between the Escherichia coli ATP-dependent proteases, ClpXP and ClpAP." J Biol Chem **273**(20): 12476-81.
Guzman, L. M., D. Belin, et al. (1995). "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter." J Bacteriol **177**(14): 4121-30.
Halbig, D., T. Wiegert, et al. (1999). "The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding." Eur J Biochem **263**(2): 543-51.
Harvey, B. R., G. Georgiou, et al. (2004). "Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries." Proc Natl Acad Sci U S A **101**(25): 9193-8.
Hatzixanthis, K., T. A. Clarke, et al. (2005). "Signal peptide-chaperone interactions on the twin-arginine protein transport pathway." Proc Natl Acad Sci U S A **102**(24): 8460-5.
Hayhurst, A. and G. Georgiou (2001). "High-throughput antibody isolation." Curr Opin Chem Biol **5**(6): 683-9.
Hayhurst, A., S. Happe, et al. (2003). "Isolation and expression of recombinant antibody fragments to the biological warfare pathogen Brucella melitensis." J Immunol Methods **276**(1-2): 185-96.
Herendeen, S. L., R. A. VanBogelen, et al. (1979). "Levels of major proteins of Escherichia coli during growth at different temperatures." J Bacteriol **139**(1): 185-94.
Hersch, G. L., T. A. Baker, et al. (2004). "SspB delivery of substrates for ClpXP proteolysis probed by the design of improved degradation tags." Proc Natl Acad Sci U S A **101**(33): 12136-41.
Herzenberg, L. A., D. Parks, et al. (2002). "The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford." Clin Chem **48**(10): 1819-27.
Hinnerwisch, J., W. A. Fenton, et al. (2005). "Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation." Cell **121**(7): 1029-41.
Hinnerwisch, J., B. G. Reid, et al. (2005). "Roles of the N-domains of the ClpA unfoldase in binding substrate proteins and in stable complex formation with the ClpP protease." J Biol Chem **280**(49): 40838-44.
Holmgren, A., C. Johansson, et al. (2005). "Thiol redox control via thioredoxin and glutaredoxin systems." Biochem Soc Trans **33**(Pt 6): 1375-7.
Holt, L. J., C. Herring, et al. (2003). "Domain antibodies: proteins for therapy." Trends Biotechnol **21**(11): 484-90.
Hoskins, J. R., S. Y. Kim, et al. (2000). "Substrate recognition by the ClpA chaperone component of ClpAP protease." J Biol Chem **275**(45): 35361-7.

Hoskins, J. R., S. K. Singh, et al. (2000). "Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP." Proc Natl Acad Sci U S A **97**(16): 8892-7.
Hoskins, J. R. and S. Wickner (2006). "Two peptide sequences can function cooperatively to facilitate binding and unfolding by ClpA and degradation by ClpAP." Proc Natl Acad Sci U S A **103**(4): 909-14.
Hoskins, J. R., K. Yanagihara, et al. (2002). "ClpAP and ClpXP degrade proteins with tags located in the interior of the primary sequence." Proc Natl Acad Sci U S A **99**(17): 11037-42.
Hoyer, W., K. Ramm, et al. (2002). "A kinetic trap is an intrinsic feature in the folding pathway of single-chain Fv fragments." Biophys Chem **96**(2-3): 273-84.
Huber, D., M. I. Cha, et al. (2005). "A selection for mutants that interfere with folding of Escherichia coli thioredoxin-1 in vivo." Proc Natl Acad Sci U S A **102**(52): 18872-7.
Huston, J. S., D. Levinson, et al. (1988). "Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli." Proc Natl Acad Sci U S A **85**(16): 5879-83.
Hutcheon, G. W. and A. Bolhuis (2003). "The archaeal twin-arginine translocation pathway." Biochem Soc Trans **31**(Pt 3): 686-9.
Hwang, B. J., W. J. Park, et al. (1987). "Escherichia coli contains a soluble ATP-dependent protease (Ti) distinct from protease La." Proc Natl Acad Sci U S A **84**(16): 5550-4.
Hynds, P. J., D. Robinson, et al. (1998). "The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane." J Biol Chem **273**(52): 34868-74.
Jack, R. L., G. Buchanan, et al. (2004). "Coordinating assembly and export of complex bacterial proteins." Embo J **23**(20): 3962-72.
Jack, R. L., F. Sargent, et al. (2001). "Constitutive expression of Escherichia coli tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth." J Bacteriol **183**(5): 1801-4.
Jager, M., P. Gehrig, et al. (2001). "The scFv fragment of the antibody hu4D5-8: evidence for early premature domain interaction in refolding." J Mol Biol **305**(5): 1111-29.
Jager, M. and A. Pluckthun (1999). "Folding and assembly of an antibody Fv fragment, a heterodimer stabilized by antigen." J Mol Biol **285**(5): 2005-19.
Jeffrey, P. D., R. K. Strong, et al. (1993). "26-10 Fab-digoxin complex: affinity and specificity due to surface complementarity." Proc Natl Acad Sci U S A **90**(21): 10310-4.
Jenal, U. and T. Fuchs (1998). "An essential protease involved in bacterial cell-cycle control." Embo J **17**(19): 5658-69.
Jenal, U. and R. Hengge-Aronis (2003). "Regulation by proteolysis in bacterial cells." Curr Opin Microbiol **6**(2): 163-72.
Joshi, S. A., G. L. Hersch, et al. (2004). "Communication between ClpX and ClpP during substrate processing and degradation." Nat Struct Mol Biol **11**(5): 404-11.
Karzai, A. W., E. D. Roche, et al. (2000). "The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue." Nat Struct Biol **7**(6): 449-55.

Katayama, Y., S. Gottesman, et al. (1988). "The two-component, ATP-dependent Clp protease of Escherichia coli. Purification, cloning, and mutational analysis of the ATP-binding component." J Biol Chem **263**(29): 15226-36.
Keiler, K. C., P. R. Waller, et al. (1996). "Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA." Science **271**(5251): 990-3.
Kenniston, J. A., T. A. Baker, et al. (2003). "Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine." Cell **114**(4): 511-20.
Kenniston, J. A., T. A. Baker, et al. (2005). "Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing." Proc Natl Acad Sci U S A **102**(5): 1390-5.
Kenniston, J. A., R. E. Burton, et al. (2004). "Effects of local protein stability and the geometric position of the substrate degradation tag on the efficiency of ClpXP denaturation and degradation." J Struct Biol **146**(1-2): 130-40.
Ki, J. J., Y. Kawarasaki, et al. (2004). "A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis." J Mol Biol **341**(4): 901-9.
Kim, D. Y. and K. K. Kim (2003). "Crystal structure of ClpX molecular chaperone from Helicobacter pylori." J Biol Chem **278**(50): 50664-70.
Kim, J. and D. A. Kendall (2000). "Sec-dependent protein export and the involvement of the molecular chaperone SecB." Cell Stress Chaperones **5**(4): 267-75.
Kim, J. Y., E. A. Fogarty, et al. (2005). "Twin-arginine translocation of active human tissue plasminogen activator in Escherichia coli." Appl Environ Microbiol **71**(12): 8451-9.
Kim, Y. I., I. Levchenko, et al. (2001). "Molecular determinants of complex formation between Clp/Hsp100 ATPases and the ClpP peptidase." Nat Struct Biol **8**(3): 230-3.
King, P. W. and A. E. Przybyla (1999). "Response of hya expression to external pH in Escherichia coli." J Bacteriol **181**(17): 5250-6.
Kirk, O., T. V. Borchert, et al. (2002). "Industrial enzyme applications." Curr Opin Biotechnol **13**(4): 345-51.
Kleerebezem, M., W. Crielaard, et al. (1996). "Involvement of stress protein PspA (phage shock protein A) of Escherichia coli in maintenance of the protonmotive force under stress conditions." Embo J **15**(1): 162-71.
Kleerebezem, M. and J. Tommassen (1993). "Expression of the pspA gene stimulates efficient protein export in Escherichia coli." Mol Microbiol **7**(6): 947-56.
Knoblauch, N. T., S. Rudiger, et al. (1999). "Substrate specificity of the SecB chaperone." J Biol Chem **274**(48): 34219-25.
Kurokawa, Y., H. Yanagi, et al. (2001). "Overproduction of bacterial protein disulfide isomerase (DsbC) and its modulator (DsbD) markedly enhances periplasmic production of human nerve growth factor in Escherichia coli." J Biol Chem **276**(17): 14393-9.
Kwon, A. R., C. B. Trame, et al. (2004). "Kinetics of protein substrate degradation by HslUV." J Struct Biol **146**(1-2): 141-7.

Leary, J. F. (1994). "Strategies for rare cell detection and isolation." Methods Cell Biol **42 Pt B**: 331-58.
Lee, C., M. P. Schwartz, et al. (2001). "ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal." Mol Cell **7**(3): 627-37.
Lee, P. A., G. Buchanan, et al. (2002). "Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation." J Bacteriol **184**(21): 5871-9.
Lee, P. A., D. Tullman-Ercek, et al. (2006). "The Bacterial Twin-Arginine Translocation Pathway." Annu Rev Microbiol.
Leipold, R. J. and P. Dhurjati (1993). "Specialized ribosomes in Escherichia coli." Biotechnol Prog **9**(5): 443-9.
Levchenko, I., R. A. Grant, et al. (2003). "Structure of a delivery protein for an AAA+ protease in complex with a peptide degradation tag." Mol Cell **12**(2): 365-72.
Levchenko, I., M. Seidel, et al. (2000). "A specificity-enhancing factor for the ClpXP degradation machine." Science **289**(5488): 2354-6.
Levy, R., R. Weiss, et al. (2001). "Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones." Protein Expr Purif **23**(2): 338-47.
Li, S. Y., B. Y. Chang, et al. (2006). "Coexpression of TorD enhances the transport of GFP via the TAT pathway." J Biotechnol **122**(4): 412-21.
Lodish, H. B., Arnold; Zipursky, S. Lawrence; Matsudaira, Paul; Baltimore, David; Darnell, James E. (2000). The Proton-Motive Force in Mitochondria Is Due Largely to a Voltage Gradient across the Inner Membrane.
Loll, P. J. (2003). "Membrane protein structural biology: the high throughput challenge." J Struct Biol **142**(1): 144-53.
Long, M., S. Lu, et al. (2006). "Kinetics of receptor-ligand interactions in immune responses." Cell Mol Immunol **3**(2): 79-86.
Makhatadze, G. I., V. V. Loladze, et al. (2003). "Contribution of surface salt bridges to protein stability: guidelines for protein engineering." J Mol Biol **327**(5): 1135-48.
Manheim, B. S., Jr., P. Granahan, et al. (2006). "'Follow-on biologics': ensuring continued innovation in the biotechnology industry." Health Aff (Millwood) **25**(2): 394-404.
Mansfeld, J., G. Vriend, et al. (1997). "Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond." J Biol Chem **272**(17): 11152-6.
Marshall, S. A., C. S. Morgan, et al. (2002). "Electrostatics significantly affect the stability of designed homeodomain variants." J Mol Biol **316**(1): 189-99.
Martin, A. (1996). "Accessing the Kabat Antibody Sequence Database by Computer." PROTEINS: Structure, Function and Genetics **25**: 130-133.
Martin, A., T. A. Baker, et al. (2005). "Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines." Nature **437**(7062): 1115-20.
Martineau, P. and J. M. Betton (1999). "In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm." J Mol Biol **292**(4): 921-9.
Martineau, P., P. Jones, et al. (1998). "Expression of an antibody fragment at high levels in the bacterial cytoplasm." J Mol Biol **280**(1): 117-27.

Martinez, M. B., M. C. Flickinger, et al. (1999). "Steady-state enzyme kinetics in the Escherichia coli periplasm: a model of a whole cell biocatalyst." J Biotechnol **71**(1-3): 59-66.
Masip, L., J. L. Pan, et al. (2004). "An engineered pathway for the formation of protein disulfide bonds." Science **303**(5661): 1185-9.
Mastrobattista, E., V. Taly, et al. (2005). "High-throughput screening of enzyme libraries: in vitro evolution of a beta-galactosidase by fluorescence-activated sorting of double emulsions." Chem Biol **12**(12): 1291-300.
Matouschek, A. (2003). "Protein unfolding--an important process in vivo?" Curr Opin Struct Biol **13**(1): 98-109.
Mattanovich, D. and N. Borth (2006). "Applications of cell sorting in biotechnology." Microb Cell Fact **5**: 12.
Maurizi, M. R., W. P. Clark, et al. (1990). "Clp P represents a unique family of serine proteases." J Biol Chem **265**(21): 12546-52.
Maurizi, M. R., S. K. Singh, et al. (1998). "Molecular properties of ClpAP protease of Escherichia coli: ATP-dependent association of ClpA and clpP." Biochemistry **37**(21): 7778-86.
McGettrick, A. F. and D. M. Worrall (2003). Extraction of Recombinant Protein From Bacteria.
McGinness, K. E., T. A. Baker, et al. (2006). "Engineering controllable protein degradation." Mol Cell **22**(5): 701-7.
Meyer-Rosberg, K., D. R. Scott, et al. (1996). "The effect of environmental pH on the proton motive force of Helicobacter pylori." Gastroenterology **111**(4): 886-900.
Mikhaleva, N. I., C. L. Santini, et al. (1999). "Requirement for phospholipids of the translocation of the trimethylamine N-oxide reductase through the Tat pathway in Escherichia coli." FEBS Lett **463**(3): 331-5.
Model, P., G. Jovanovic, et al. (1997). "The Escherichia coli phage-shock-protein (psp) operon." Mol Microbiol **24**(2): 255-61.
Mori, H. and K. Cline (2002). "A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase." J Cell Biol **157**(2): 205-10.
Mori, H. and K. Ito (2001). "The Sec protein-translocation pathway." Trends Microbiol **9**(10): 494-500.
Murphy, K. and A. Boyet. (2004). "FACS Lab." from http://www.pathology.wustl.edu/html/facilitiesfacs.html.
Musser, S. M. and S. M. Theg (2000). "Proton transfer limits protein translocation rate by the thylakoid DeltapH/Tat machinery." Biochemistry **39**(28): 8228-33.
Nakonechny, W. S. and C. M. Teschke (1998). "GroEL and GroES control of substrate flux in the in vivo folding pathway of phage P22 coat protein." J Biol Chem **273**(42): 27236-44.
Neidhardt, F. C. (1996). *Escherichia coli* and *Salmonella*. Washington, ASM Press.
Oates, J., J. Mathers, et al. (2003). "Consensus structural features of purified bacterial TatABC complexes." J Mol Biol **330**(2): 277-86.
Ogura, T. and A. J. Wilkinson (2001). "AAA+ superfamily ATPases: common structure--diverse function." Genes Cells **6**(7): 575-97.

Olsen, M., B. Iverson, et al. (2000). "High-throughput screening of enzyme libraries." Curr Opin Biotechnol **11**(4): 331-7.
Olsen, M. J., J. Gam, et al. (2003). "High-throughput FACS method for directed evolution of substrate specificity." Methods Mol Biol **230**: 329-42.
Orlandi, R., D. H. Gussow, et al. (1989). "Cloning immunoglobulin variable domains for expression by the polymerase chain reaction." Proc Natl Acad Sci U S A **86**(10): 3833-7.
Ortega, J., H. S. Lee, et al. (2002). "Alternating translocation of protein substrates from both ends of ClpXP protease." Embo J **21**(18): 4938-49.
Ortega, J., H. S. Lee, et al. (2004). "ClpA and ClpX ATPases bind simultaneously to opposite ends of ClpP peptidase to form active hybrid complexes." J Struct Biol **146**(1-2): 217-26.
Pace, C. N. (1990). "Measuring and increasing protein stability." Trends Biotechnol **8**(4): 93-8.
Pak, M., J. R. Hoskins, et al. (1999). "Concurrent chaperone and protease activities of ClpAP and the requirement for the N-terminal ClpA ATP binding site for chaperone activity." J Biol Chem **274**(27): 19316-22.
Palmer, T. and B. C. Berks (2003). "Moving folded proteins across the bacterial cell membrane." Microbiology **149**(Pt 3): 547-56.
Palmer, T., F. Sargent, et al. (2005). "Export of complex cofactor-containing proteins by the bacterial Tat pathway." Trends Microbiol **13**(4): 175-80.
Paul, R., J. Apel, et al. (2003). "Shear stress related blood damage in laminar couette flow." Artif Organs **27**(6): 517-29.
Perez, O. D. and G. P. Nolan (2002). "Simultaneous measurement of multiple active kinase states using polychromatic flow cytometry." Nat Biotechnol **20**(2): 155-62.
Pohlschroder, M., M. I. Gimenez, et al. (2005). "Protein transport in Archaea: Sec and twin arginine translocation pathways." Curr Opin Microbiol **8**(6): 713-9.
Pommier, J., V. Mejean, et al. (1998). "TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in Escherichia coli." J Biol Chem **273**(26): 16615-20.
Porankiewicz, J., J. Wang, et al. (1999). "New insights into the ATP-dependent Clp protease: Escherichia coli and beyond." Mol Microbiol **32**(3): 449-58.
Presta, L. G. (2006). "Engineering of therapeutic antibodies to minimize immunogenicity and optimize function." Adv Drug Deliv Rev.
Puchkaev, A. V., L. S. Koo, et al. (2003). "Aromatic stacking as a determinant of the thermal stability of CYP119 from Sulfolobus solfataricus." Arch Biochem Biophys **409**(1): 52-8.
Qiu, J., J. R. Swartz, et al. (1998). "Expression of active human tissue-type plasminogen activator in Escherichia coli." Appl Environ Microbiol **64**(12): 4891-6.
Raag, R. and M. Whitlow (1995). "Single-chain Fvs." Faseb J **9**(1): 73-80.
Randall, L. L., J. M. Crane, et al. (2004). "Sites of interaction between SecA and the chaperone SecB, two proteins involved in export." Protein Sci **13**(4): 1124-33.
Randall, L. L. and S. J. Hardy (2002). "SecB, one small chaperone in the complex milieu of the cell." Cell Mol Life Sci **59**(10): 1617-23.
Reichert, J. M., C. J. Rosensweig, et al. (2005). "Monoclonal antibody successes in the clinic." Nat Biotechnol **23**(9): 1073-8.

Richter, S. and T. Bruser (2005). "Targeting of unfolded PhoA to the TAT translocon of Escherichia coli." J Biol Chem **280**(52): 42723-30.
Robinson, C. and A. Bolhuis (2001). "Protein targeting by the twin-arginine translocation pathway." Nat Rev Mol Cell Biol **2**(5): 350-6.
Robinson, C., C. Woolhead, et al. (2000). "Transport of proteins into and across the thylakoid membrane." J Exp Bot **51 Spec No**: 369-74.
Roche, E. D. and R. T. Sauer (1999). "SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity." Embo J **18**(16): 4579-89.
Rodrigue, A., A. Chanal, et al. (1999). "Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway." J Biol Chem **274**(19): 13223-8.
Rose, R. W., T. Bruser, et al. (2002). "Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway." Mol Microbiol **45**(4): 943-50.
Schatz, G. and B. Dobberstein (1996). "Common principles of protein translocation across membranes." Science **271**(5255): 1519-26.
Schierle, C. F., M. Berkmen, et al. (2003). "The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway." J Bacteriol **185**(19): 5706-13.
Schwan, W. R., J. L. Lee, et al. (2002). "Osmolarity and pH growth conditions regulate fim gene transcription and type 1 pilus expression in uropathogenic Escherichia coli." Infect Immun **70**(3): 1391-402.
Segatori, L., P. J. Paukstelis, et al. (2004). "Engineered DsbC chimeras catalyze both protein oxidation and disulfide-bond isomerization in Escherichia coli: Reconciling two competing pathways." Proc Natl Acad Sci U S A **101**(27): 10018-23.
Sharkey, R. M. and D. M. Goldenberg (2006). "Targeted therapy of cancer: new prospects for antibodies and immunoconjugates." CA Cancer J Clin **56**(4): 226-43.
Sharma, S., J. R. Hoskins, et al. (2005). "Binding and degradation of heterodimeric substrates by ClpAP and ClpXP." J Biol Chem **280**(7): 5449-55.
Singh, S. K., R. Grimaud, et al. (2000). "Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP." Proc Natl Acad Sci U S A **97**(16): 8898-903.
Singh, S. K., J. Rozycki, et al. (2001). "Functional domains of the ClpA and ClpX molecular chaperones identified by limited proteolysis and deletion analysis." J Biol Chem **276**(31): 29420-9.
Singleton, M. R., M. R. Sawaya, et al. (2000). "Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides." Cell **101**(6): 589-600.
Smith, C. K., T. A. Baker, et al. (1999). "Lon and Clp family proteases and chaperones share homologous substrate-recognition domains." Proc Natl Acad Sci U S A **96**(12): 6678-82.
Song, H. K. and M. J. Eck (2003). "Structural basis of degradation signal recognition by SspB, a specificity-enhancing factor for the ClpXP proteolytic machine." Mol Cell **12**(1): 75-86.

Stanley, N. R., K. Findlay, et al. (2001). "Escherichia coli strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope." J Bacteriol **183**(1): 139-44.
Stanley, N. R., T. Palmer, et al. (2000). "The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli." J Biol Chem **275**(16): 11591-6.
Stanley, N. R., F. Sargent, et al. (2002). "Behaviour of topological marker proteins targeted to the Tat protein transport pathway." Mol Microbiol **43**(4): 1005-21.
Straus, D. B., W. A. Walter, et al. (1988). "Escherichia coli heat shock gene mutants are defective in proteolysis." Genes Dev **2**(12B): 1851-8.
Swartz, J. R. (2001). "Advances in Escherichia coli production of therapeutic proteins." Curr Opin Biotechnol **12**(2): 195-201.
Szyk, A. and M. R. Maurizi (2006). "Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site." J Struct Biol.
Thibault, G., Y. Tsitrin, et al. (2006). "Large nucleotide-dependent movement of the N-terminal domain of the ClpX chaperone." Embo J **25**(14): 3367-76.
Thomas, J. D., R. A. Daniel, et al. (2001). "Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli." Mol Microbiol **39**(1): 47-53.
Thomas, J. G., A. Ayling, et al. (1997). "Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. To fold or to refold." Appl Biochem Biotechnol **66**(3): 197-238.
Thomas, J. G. and F. Baneyx (1996). "Protein folding in the cytoplasm of Escherichia coli: requirements for the DnaK-DnaJ-GrpE and GroEL-GroES molecular chaperone machines." Mol Microbiol **21**(6): 1185-96.
Thompson, M. W. and M. R. Maurizi (1994). "Activity and specificity of Escherichia coli ClpAP protease in cleaving model peptide substrates." J Biol Chem **269**(27): 18201-8.
Thompson, M. W., S. K. Singh, et al. (1994). "Processive degradation of proteins by the ATP-dependent Clp protease from Escherichia coli. Requirement for the multiple array of active sites in ClpP but not ATP hydrolysis." J Biol Chem **269**(27): 18209-15.
Thomsen, L. E., J. E. Olsen, et al. (2002). "ClpP is involved in the stress response and degradation of misfolded proteins in Salmonella enterica serovar Typhimurium." Microbiology **148**(Pt 9): 2727-33.
Tinker, J. K., J. L. Erbe, et al. (2005). "Characterization of fluorescent chimeras of cholera toxin and Escherichia coli heat-labile enterotoxins produced by use of the twin arginine translocation system." Infect Immun **73**(6): 3627-35.
Tsien, R. Y. (1998). "The green fluorescent protein." Annu Rev Biochem **67**: 509-44.
Tyo, K. E., H. Zhou, et al. (2006). "High-throughput screen for poly-3-hydroxybutyrate in Escherichia coli and Synechocystis sp. strain PCC6803." Appl Environ Microbiol **72**(5): 3412-7.
Uchida, K., H. Mori, et al. (1995). "Stepwise movement of preproteins in the process of translocation across the cytoplasmic membrane of Escherichia coli." J Biol Chem **270**(52): 30862-8.

Veenendaal, A. K., C. van der Does, et al. (2004). "The protein-conducting channel SecYEG." Biochim Biophys Acta **1694**(1-3): 81-95.
Villaverde, A. and M. M. Carrio (2003). "Protein aggregation in recombinant bacteria: biological role of inclusion bodies." Biotechnol Lett **25**(17): 1385-95.
Vitetta, E. S. and V. F. Ghetie (2006). "Immunology. Considering therapeutic antibodies." Science **313**(5785): 308-9.
Voges, D., P. Zwickl, et al. (1999). "The 26S proteasome: a molecular machine designed for controlled proteolysis." Annu Rev Biochem **68**: 1015-68.
Wah, D. A., I. Levchenko, et al. (2002). "Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer." Chem Biol **9**(11): 1237-45.
Walsh, G. (2005). "Biopharmaceuticals: recent approvals and likely directions." Trends Biotechnol **23**(11): 553-8.
Wang, J., J. A. Hartling, et al. (1997). "The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis." Cell **91**(4): 447-56.
Wang, J., J. J. Song, et al. (2001). "Nucleotide-dependent conformational changes in a protease-associated ATPase HsIU." Structure **9**(11): 1107-16.
Wexler, M., F. Sargent, et al. (2000). "TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export." J Biol Chem **275**(22): 16717-22.
Wickner, S., S. Gottesman, et al. (1994). "A molecular chaperone, ClpA, functions like DnaK and DnaJ." Proc Natl Acad Sci U S A **91**(25): 12218-22.
Wickner, W. and R. Schekman (2005). "Protein translocation across biological membranes." Science **310**(5753): 1452-6.
Wulfing, C. and A. Pluckthun (1994). "Protein folding in the periplasm of Escherichia coli." Mol Microbiol **12**(5): 685-92.
Wyborn, N. R., S. L. Messenger, et al. (2002). "Expression of the Escherichia coli yfiD gene responds to intracellular pH and reduces the accumulation of acidic metabolic end products." Microbiology **148**(Pt 4): 1015-26.
Xia, D., L. Esser, et al. (2004). "Crystallographic investigation of peptide binding sites in the N-domain of the ClpA chaperone." J Struct Biol **146**(1-2): 166-79.
Yahr, T. L. and W. T. Wickner (2001). "Functional reconstitution of bacterial Tat translocation in vitro." Embo J **20**(10): 2472-9.
Zellmeier, S., W. Schumann, et al. (2006). "Involvement of Clp protease activity in modulating the Bacillus subtilissigmaw stress response." Mol Microbiol **61**(6): 1569-82.
Zhang, Z., Z. H. Li, et al. (2002). "Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in Escherichia coli." Protein Expr Purif **26**(2): 218-28.
Zimmer, M. (2002). "Green fluorescent protein (GFP): applications, structure, and related photophysical behavior." Chem Rev **102**(3): 759-81.

Vita

Brian Michael Ribnicky was born in Morristown, New Jersey on March 23, 1975, the son of Anastasia Ribnicky and Michael John Ribnicky. He grew up in the township of Roxbury, New Jersey where he graduated from Roxbury High School in 1993. He then entered Stevens Institute of Technology in Hoboken, New Jersey, where he graduated in 1997 as valedictorian with the degree of Bachelor of Engineering in Chemical Engineering with minors in Social Science and Biochemical Engineering. Brian then accepted a chemical engineering position with BASF Corporation and worked at several site locations in the eastern half of the United States. In June 1999, he left BASF and joined Merck & Co. as a manufacturing engineer at their Flint River bulk pharmaceutical production site in Albany, GA. Wanting to further his education, in August 2001, he entered the Graduate School of The University of Texas at Austin to pursue his Doctorate in Chemical Engineering. In December 2003, Brian received the degree of Master of Engineering in Chemical Engineering. He has one brother, a graduate of Saint Peter's College employed in new home construction management.

Permanent address: 324 Reynolds Rd, Egg Harbor Township, New Jersey 08234

This dissertation was typed by the author.